Vol. 37, No. 4 April, 1940

Psychological Bulletin

CURRENT ISSUES IN FACTOR ANALYSIS[1]

BY L. L. THURSTONE
University of Chicago

THE PURPOSE OF FACTOR ANALYSIS

A factor problem starts with the hope or conviction that a certain domain is not so chaotic as it looks. The object of factor analysis is to discover the principal dimensions or categories of mentality and to indicate the directions along which they may be studied by experimental laboratory methods. Factor analysis is not restricted by assumptions regarding the nature of the factors, whether they be physiological or social, elemental or complex, correlated or uncorrelated. It assumes that a variety of phenomena within the domain are related and that they are determined, at least in part, by a relatively small number of functional unities, or factors. The factors may be called by different names, such as causes, faculties, parameters, functional unities, abilities, independent measurements, experimentally independent effects. The name for a factor depends on the context, on one's philosophical preferences and manner of speech, and on how much one already knows about the domain to be investigated. The factors in psychological investigations are not ordinarily to be thought of as elemental things which are present or absent, like heads or tails in the tossing of coins.

On previous occasions I have stated that factor analysis has its principal usefulness at the borderline of science and that it is naturally superseded as quickly as possible by rational formulations in terms of the science involved. Factor analysis is useful especially in

[1] The statistical work on various phases of factor analysis described in this paper was done with the assistance of research grants from the Carnegie Corporation of New York and from the Social Science Research Committee at the University of Chicago.

those domains where basic and fruitful concepts are essentially lacking and where crucial experiments have been difficult to conceive. The new methods have a humble role. They enable us to make only the crudest first map of a new domain. But if we have scientific intuition and sufficient ingenuity, the rough factorial map of a new domain will enable us to proceed beyond the factorial stage to the more direct forms of psychological experimentation in the laboratory. I fear that this exploratory nature of factor analysis is often not understood.

In factorial investigations we proceed on the assumption that mind is structured somehow, that mind is not a patternless mosaic of an infinite number of elements without functional groupings. In speaking of a man's muscular development we can legitimately say that his arm is strong. Arm strength could appear as a factor in a battery of tests of muscular development. But if we should find such a factor as arm strength, we could speak of the factor without assuming that it was an ultimate indivisible element.

Observation and educational experience lend plausibility to the conception that the mental abilities are determined by a great multiplicity of causes or determiners and that these determiners are more or less structured or linked in groups. This multiplicity of determiners can be thought of as a field of elements in which all are not equally closely linked. Some elements may be quite independent in their action, while others may be rather closely associated. The factors are probably functional groupings, and it is a distortion to assume that they must be elemental. We know precious little about the determiners of human talent and temperament, and we should not impose on our thinking an unnecessarily rigid causal frame.

If we grant that men are not all equal in intellectual endowment and in temperament, and if we have the faith that this domain can be investigated as science, then we must make the plausible and inevitable assumption that individual differences among men can be conceived in terms of a definite number of traits, parameters, or factors. Some of the factors may be found to be anatomically determined; others will be physiological; still others will be defined at first in experiential, educational, and social terms. As scientists we must believe that a set of categories can be found for the understanding of mentality, which have, by their simplicity, a prior claim on our conceptual formulations. This reasoning seems rather too obvious to be debatable.

Simple Structure

The concept of simple structure, or simple configuration, can be explained psychologically and entirely apart from the mathematical form of the solution which the psychological concept determines. Such an explanation will be attempted here.

In Figure 1 the large square represents a table of intercorrelations, known as a *correlation matrix*. The factor problem starts ordinarily with such a table. The object of the factor analysis is to find a corresponding *factor matrix F,* shown at the right in the figure. This factor matrix must have a number of columns much smaller

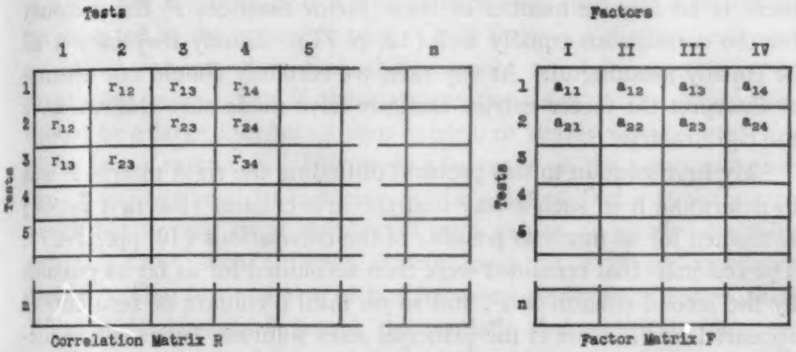

FIGURE 1

than the number of columns in the correlation matrix if the result is to be useful. The number of columns in F is the number of factors that must be postulated in order to account for the correlations in R. Now, if there are as many columns in the factor matrix F as there are tests in the correlation matrix R, then we have as many factors as there are tests. In other words, we should have to assume a new factor for every new test in the battery. Such an arrangement would serve no economy of thought, and it would reveal no underlying groupings or simplification in the test battery. It would be of no psychological interest. We demand, in order for the factor analysis to be scientifically significant, that the number of columns in the factor matrix F must be much smaller than the number of columns in the correlation matrix R. This condition is not alone sufficient to make the analysis of scientific interest.

Suppose that we have found a factor matrix F with a small number of columns that accounts for the correlations in R. If the matrix F is to be scientifically significant, we should be able to read meaning

in the numerical entries in F. In the diagram we have shown four columns in the factor matrix as an example. The numbers in the first row of F should tell us how much of each factor is involved in the first test, and so on for each test. If the first test involves mostly factor III, then the third entry in the first row of F should be large, and the other numbers in that row should be small. In this way it should be possible to tell by inspection which of the four factors are involved in each of the tests. If we take the first column of F and locate the highest numbers in it, they would be in the rows whose tests have a large saturation in that factor.

The difficulty at this point was at first a serious one, namely, that there is an infinite number of these factor matrices F that account for the correlations equally well (**12**, p. 71). Surely they cannot all be equally meaningful. At any rate, we certainly should not attempt to interpret the factor entries until we have made sure that we have the right matrix F.

My first solution to this problem of finding the right matrix F was to determine it in such a way that the first column (the first factor) accounted for as much as possible of the correlations (**10**, pp. 17–27). The residuals that remained were then accounted for as far as possible by the second column in F, and so on until a column of zero entries appeared in F. This is the principal axes solution, but it was inadequate because of a psychological consideration. The factorial composition of any particular test A could be altered by merely inserting test A in different test batteries, and this did not make sense. The primary abilities that are required for a certain test A cannot be influenced by the other tests that the subjects might be asked to take after completing test A. The numerical entries in a row of the matrix F that describe a particular test would depend not only on the abilities involved in test A, but also on the rest of the tests with which A might be arbitrarily combined. If there are such things as distinct mental abilities and traits which enter into the test performances, then these should be represented in the factor matrix in some unique way. The centroid method of factoring the correlation matrix has the same defect, in that the reference frame of F must be rotated into a unique and meaningful position before interpretation can be attempted with any hope of success. This is true for all methods of factoring that have been devised so far.

The problem of finding the right factorial matrix out of an infinite number of possible factor matrices remained a stumbling block until I found the principle that I have called simple structure (**12**, Chap.

VI). It was arrived at by psychological considerations and not by any statistical reasoning. Suppose that there are distinct mental abilities and traits, such as facility with numbers, ability in rhyming, a good memory, ability to think of objects in solid space, quickness in perceiving visual detail, auditory acuity, ease in gaining a large vocabulary, verbal fluency, ability to do crossword puzzles, and so on. Let the list be whatever the unknown human abilities are. Consider now any of the tests, such as naming the opposite of words, column addition, or simple visualizing problems. It seems psychologically reasonable that each of these tests will not require all of the abilities that are required by the test battery as a whole. If there is a distinct ability to deal with numbers, then that factor would not be expected in the opposites test. Or if there is a distinct trait such as verbal fluency, then that ability would not be expected in the visualizing tasks. But if these expectations are correct, then there should be a large number of zero entries, or nearly vanishing entries, in the factor matrix F. This gives a leverage on the rotational problem of finding the right factorial matrix. Let us find a factor matrix which has a very large number of zero values or nearly vanishing entries. Does such a factor matrix exist for a given correlation matrix? That is a question of fact for any given table of correlations. When such a matrix is found, the result is that each individual test can be accounted for by fewer factors than are required for the whole test battery; and this seems to be psychologically reasonable.

In general, the factor matrix as obtained by different methods has about as many negative values as positive values. There is difficulty in the interpretation of negative entries in the factor matrix. A negative entry in F would have to be interpreted to mean that the possession of an ability has a detrimental effect on the test performance. One can readily understand how the possession of a certain ability can aid in a test performance, and one can imagine that an ability has no effect on a test performance, but it is difficult to think of abilities that are as often detrimental as helpful in the test performances. Surely, the correct factor matrix for cognitive tests does not have many negative entries, and preferably it should have none at all. Now it so happens, over and over again, that when a factor matrix is found with a very large number of zero entries, the negative entries disappear at the same time. It does not seem as if all of this could happen by chance. The reason is probably to be found in the underlying distinct mental processes that are involved in the different tasks. In other words, the results point to the conclusion that mind

is not a structureless mass, but that it is structured into constellations or groupings of processes that can be identified as distinct functions in the test performances. These are what I have called primary mental abilities or traits. Furthermore, when the columns of the rotated matrix F are inspected for the interpretation of the factors, it is found that the factors can be identified by tentative psychological descriptions.

Some students of factor analysis do not accept the possibility of a psychologically unique factor matrix F. They point to the fact that there is an infinite number of factor matrices F which account equally well for the correlations in R, and they conclude that the factors are, therefore, essentially mathematical artifacts. In dealing with this problem I have had the conviction that one factor matrix should exist which is not an artifact but is psychologically meaningful. The fact that it is mathematically possible to write many other arbitrary sets of numbers in the factor matrix that fit the numbers in the correlation matrix merely raises the question how to find the right one. A physical illustration can be given to show the principle involved here. If a billiard ball is given a stroke in a certain direction, the ball may be assumed to roll in that direction. A conceptual formulation might begin with a diagram of the billiard table and one arrow to show the direction of the stroke. Now it is also possible to draw an infinite number of different diagrams with two or any larger number of force vectors whose resultant would make the ball roll in exactly the same direction. If we saw the ball rolling after it had been struck, we might prove mathematically that the problem is insoluble because there is an infinite number of possible combinations of forces that *could* have made the ball roll just that same way. We might be so impressed with the absolute correctness of the proof that we would fail to look for a simple solution—that perhaps the ball was struck only once.

The indeterminacy of the factor matrix consists, after all, merely in the unknown reference frame for the test vectors. Consider a simple case of such an indeterminate reference frame. If you were asked to calculate the volume of a box, you would measure the length, width, and thickness. You would take the same measurements on the box whether it were standing upright on the floor or leaning against a post. You would unconsciously adapt your reference frame to the box. But you could use any one of an infinite number of reference frames. You might take as a reference frame a vertical line

through one corner of the box and the north and east lines through the same corner. Then you could project the box faces against the reference planes. You *could* find the volume that way, but you would prefer to think about the problem in terms of the reference frame that functioned when the box was made. This is the idea of a simple structure, a reference frame which simplifies our conception of the domain that is to be investigated and comprehended. But if it should happen that the objects to be described are formless, irregular clumps, then all reference frames would be equally awkward and no simple structure would exist. It is my conviction that mentality is, in this sense, not formless, but that it is structured somehow into constellations of processes which will eventually be identified.

As far as my own work in factor analysis is concerned, one of the turning points toward a successful solution was the discovery of the principle which I have called simple structure, or simple configuration. A simple structure can be defined statistically as a factor matrix in which a very large number of zero entries appear. The number of vanishing entries must be much larger than the number of zeros which can be introduced in an arbitrary matrix by rotation of the reference frame. In practice, one should find half or more of the entries vanishing for most of the columns. So far, there has not been developed any rigorous mathematical criterion for determining the uniqueness of the factor matrix with a simple structure. The fact that several different experiments yield the same set of identifiable primary factors is a strong indication of their uniqueness. For example, in a battery of 56 tests there were five tests with significant entries in one of the columns (16, p. 115). What were those tests? All of them turned out to be memory tests. Up to this point in the factor analysis the tests had been identified by numbers, and it was gratifying to discover some kind of order in the test performances. Similar test batteries have been given to other populations of roughly comparable age range, and the same factors appear in each of them. In the face of findings like these, I am sure that primary mental abilities are functionally significant, and I am left quite unconvinced by the critics who deny their reality.

Interpretation of Factors

As I have pointed out before, the factor methods involve no assumptions whatever as to the nature of the factors. They may be

physical or psychical, native or acquired, physiological, chemical, or social in character. They may have a significance only for the particular group investigated. For example, if a popular teacher in a school has taught several subjects, the student performances might show a factor common to these subjects. The interpretation of the factor would be the instruction by that teacher, and a knowledge of local circumstances would be required for the interpretation. Of course, we avoid as far as possible such local factors in most investigations; but their exclusion can never be guaranteed beforehand in any experiment. On the other hand, such factors might actually be the object of study. In fundamental psychological studies these local factors are usually of secondary interest, and they are avoided as far as possible in factor studies as well as in most scientific work.

A combination of native and educational factors can be found in the same test battery. A class of boys might show individual differences in ability to perform geometrical tasks. One of the factors could be native, the ability to visualize space, and another factor might be the amount of instruction in mechanical drawing. The interpretation of the factors calls for psychological insight. It is not merely a statistical matter.

When the rotated factor matrix has been found, the interpretation of each factor is rather simple in principle, but it is not always easy to carry out. Each column is inspected to find a list of measures or tests which have a high saturation in the factor that is represented by the column. Another list is made for the tests in which the factor is absent or negligible. The most important part of a factor analysis is, then, the interpretation of the factors. What is it that variables in the first list have in common which is absent or negligible in the second list? If the configuration happens to be very clear, an interpretation can be expected. Sometimes several interpretations may suggest themselves. Such a situation calls for a new factor experiment to test the several hypotheses.

A few examples from our experience with factor experiments will illustrate the nature of this problem. Consider, as an illustration, the factor that I have called *N* to represent facility with numbers. In several of my factorial investigations this dimension was found, and in Figure 2 we have the distribution of test variances that are attributable to this factor in a large battery of 56 tests (**16**, p. 115). Note that only eight of the tests have an appreciable part of their variances attributable to this factor. Its functional uniqueness was

suspected by the fact that it is a dimension which claims none of the common-factor variance of most of the tests but which is conspicuous in five of the tests. Can this factor represent anything that can be given psychological interpretation? We look at the five tests and find: Addition, Subtraction, Multiplication, Division, Number Code, and, to a smaller extent, the tests called Numerical Judgment, Tabular Completion, and Arithmetical Reasoning. One should have to be

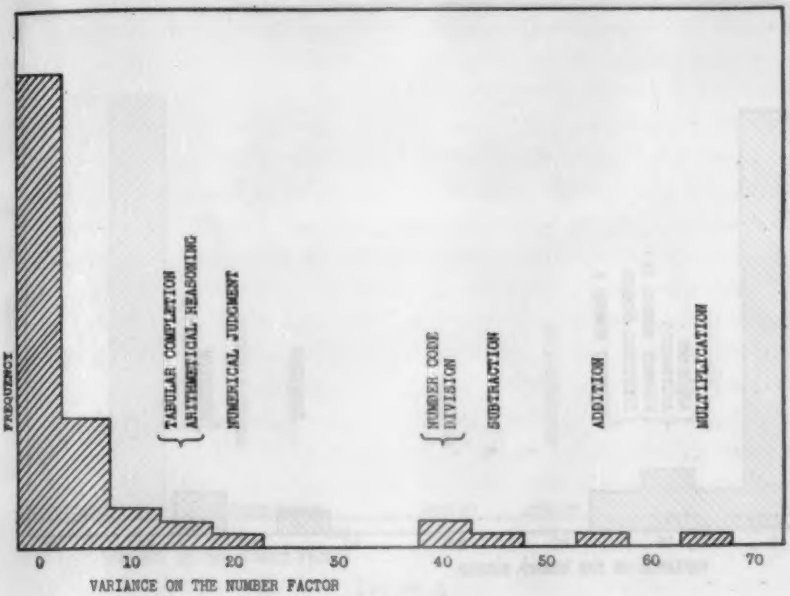

FIGURE 2

stupid indeed not to notice that these tests are numerical in character. What about all of the 48 other tests with zero variance on this factor? Are any of them numerical? If this dimension really represents numerical facility, then it must include all of the tests which obviously involve arithmetical manipulation. Inspection of the rest of the tests shows that no number test has zero variance on this common factor. All of the number tests have an appreciable part of their variance in this common factor.

A reduced battery of 36 tests was given to a less selected group of high-school seniors at the Lane Technical High School in Chicago. The correlational product-moment matrix was factored and rotated to a simple configuration (**14**). One of the factors is represented in

Figure 3. Again the tests with relatively high saturation in this factor were numerical. A third group of 300 seniors at the Hyde Park High School were given still a different set of tests (results not yet published). The same factor appeared again with the number tests high and with all other tests of zero variance on this factor, as shown in Figure 4. The number tests insist on separating themselves out on one of the common factors. Shall we ignore it? I

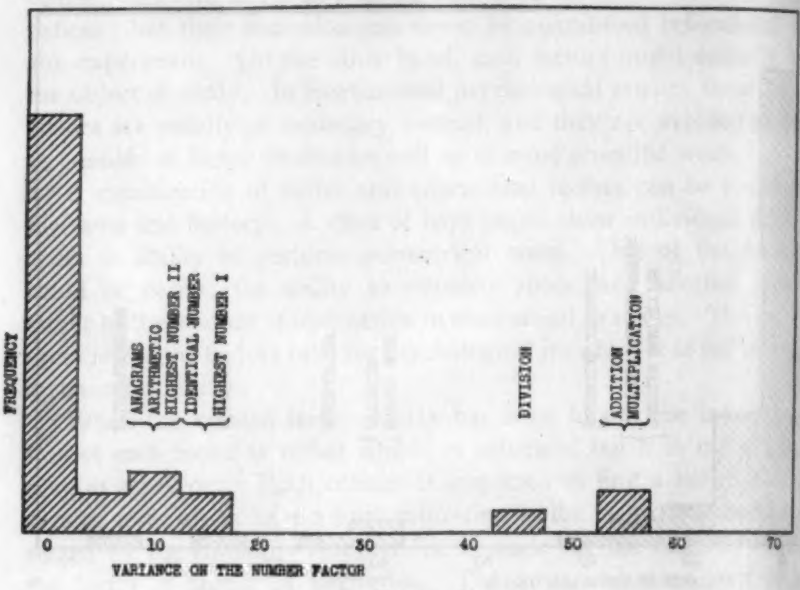

FIGURE 3

have preferred to recognize it by identifying the same factor that appears in different test batteries in different populations.

A different battery of 29 tests was given by Woodrow in connection with a study of improvement by practice (**20**). He identified his factors as numerical, perceptual, spatial, verbal, and speed; and he found an additional factor that he has tentatively interpreted as an attention factor. The first four of these factors are identical with the factors I have found with different test batteries on different populations.

When the factor has been identified as a functional unity or grouping, and given a factorial name or symbol such as *N* for number, the task of interpretation has only begun. It is an interesting psychological problem to ascertain its fundamental nature. The

factor should be identified either in the school practice or in the endowment of the subjects, or in both of these sources. If it is a schooling effect, then it should be possible to verify its nature by a test battery which includes different kinds of number tests, those which have been drilled as such, like column addition and long division, and other tests which involve manipulation with numbers in unusual tasks that are not drilled in the schools. If all of these

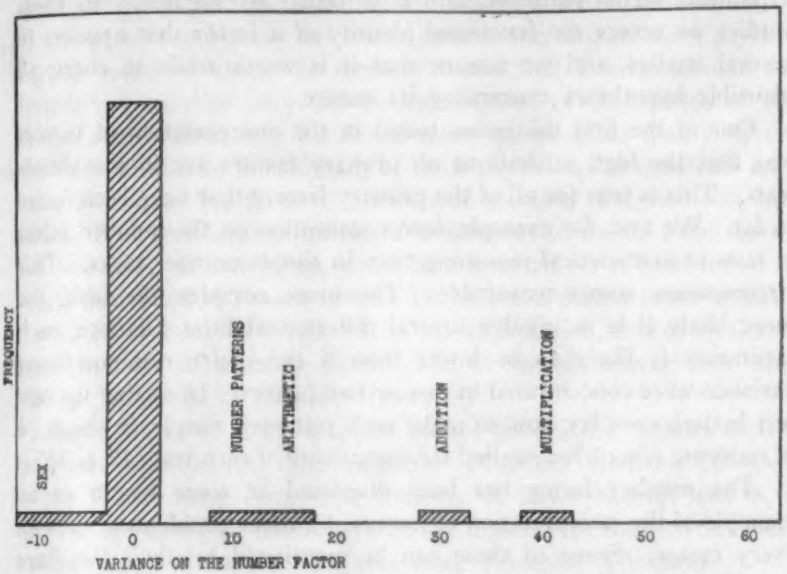

FIGURE 4

tests show high saturation on the same factor, then the interpretation is not in the formal school drill but in something more general, such as facility with the number concepts. Such a test battery is now being investigated.

There is a possibility that the factor *N* is fundamentally characterized by some dynamic aspect of simple number work which transcends the numerical nature of the tests. One student, Mr. Herbert Landahl, has made the suggestion that the number factor may be characterized by facility with a certain type of chain association and that this dynamic factor happens to be present in the number tests but not in the other tests that we have used so far. If this suggestion is correct, then it is only an historical accident that the factor *N* was discovered in tests that happen to be numerical. Further, it should be possible to test his hypothesis by a factor analysis of a specially designed test

battery. In a current investigation we have included some nonnumerical tests which involve chain association of the particular type involved in Landahl's hypothesis. It will be interesting to see if any such nonnumerical test shows high saturation on the factor that is characterized at present by the simple number tests.

Another suggestion, made by Mr. Clyde Coombs, is that the number factor is a facility with highly practiced associations. The hypothesis seems plausible, and it is being investigated. In these studies we accept the functional identity of a factor that appears in several studies, and we assume that it is worth while to check all plausible hypotheses concerning its nature.

One of the first things we found in the interpretation of factors was that the high saturations on primary factors are in the simple tests. This is true for all of the primary factors that have been found so far. We find, for example, lower saturation on the number factor in tests of arithmetical reasoning than in simple number tasks. This circumstance seems reasonable. The more complex the task, the more likely it is to involve several different abilities. Hence, each saturation is likely to be lower than if the entire common-factor variance were concentrated in one or two factors. In setting up new test batteries we try now to make each test very simple, in the hope of reducing what I have called the complexity of each test (**12**, p. 155).

The number factor has been discussed at some length as an example of the interpretation of factors. Similar problems arise with every factor. Some of these can be mentioned briefly. We have found two verbal factors. One of these is called verbal and denoted V. The other is called a word factor and denoted W. The factor V is so insistent that its existence as a functional unity or grouping is quite clear. The nature of the word factor is not so clear. It seems to be more restricted; and it may be, at least partially, independent of the meanings of the words. This factor should be explored further with atypical subjects, including reading disabilities, various articulatory disturbances, and certain types of brain lesions, in order that its fundamental nature may be clearer. It is certainly not identical with the understanding of ordinary prose.

A large test battery was recently given to 1100 eighth-grade children. The factorial analysis has just been completed. The test battery included a number of new tests which were specially constructed to test our hypothesis concerning the word-fluency factor W. According to this hypothesis, the factor W is revealed by readiness in producing words to fit a specified set. According to this hypothesis

the recognition form of an opposites test, for example, should require little or no *W*, whereas a free-writing form of the same test should have high saturation in *W*. The results show a clear separation between the factors *V* and *W* as significantly correlated, but independent, primary factors. It is my hypothesis that the factor *W* is the factor that is involved in the functional loss of amnesic aphasia. Our description of the factor *W* is remarkably similar to clinical descriptions of the functions that are lost in amnesic aphasia.

A special test battery has been assembled to answer some psychological questions about the memory factor that I have denoted *M*. Retentivity may or may not be independent of the nature of the things remembered. Is incidental memory the same ability as the ability to memorize intentionally? Such psychological questions may be answered by the factorial analysis of specially designed test batteries. The well-known individual differences in memory for names but not faces, or vice versa, could be accounted for if both forms of recall demanded the same retentivity *M* but different other abilities for visual and verbal recall. It is my expectation, however, that a small number of memory factors will be found—maybe only one, but not hundreds of memory factors.

Some critics of factor analysis do not believe that factors should be given psychological interpretation. Professor Godfrey Thomson is one of the most competent students of factor theory, and in a number of papers he has reiterated his disbelief in the psychological interpretation of factors. I quote from Professor Thomson:

"Briefly, my attitude is that I do not believe in 'factors' if any degree of real existence is attributed to them; but that of course I recognize that any set of correlated human abilities can always be described mathematically by a number of uncorrelated variables or 'factors,' *and that in many ways*" (**4**, p. 2). "My own belief is that the mind is not divided up into 'unitary factors,' but is a rich, comparatively undifferentiated complex of innumerable influences; on the physiological side an intricate network of possibilities of intercommunication" (**4**, p. 7).

"The desire to find 'realities' behind the phenomena appears to be strong in Thurstone. His conception of 'simple structure' among factors, and his belief that when 'simple structure' is achieved the factors have a significance more than that which attaches to mere statistical coefficients, is of the greatest interest" (**8**, p. 4). "I fear, however, that the factors found from different batteries by the strict application of this mathematical principle may not correspond to one another" (**8**, p. 5). "The chief deduction which can be drawn from the comparatively low rank to which so many matrices of mental correlation coefficients can be reduced is, in my opinion, the conclusion that the mind of man is com-

paratively undifferentiated, protean and plastic, *not* that it is composed of separate faculties" (8, p. 7).

"But I may, I hope, be allowed to warn against the danger of personifying these factors, which are, I think, only mathematical coefficients" (5, p. 7).

In his latest volume (9, p. 267), Professor Thomson repeats practically the same statements that we have quoted. On the question of whether mind is structured, he says: "Instead of showing that the mind has a definite structure, being composed of a few factors which work through innumerable specific machines, the low rank shows that the mind has hardly any structure" (9, p. 270). But these quotations may not adequately express his present views because of further qualifications. He finds "that a general tendency is noticeable in experimental reports to the effect that batteries do not permit of being explained by as small a number of factors in adults as in children, probably because in adults education and vocation have imposed a structure on the mind which is absent in the young"(9, p. 270). Some form of acquired structure is apparently admitted in Professor Thomson's factorial interpretations. In contrasting his sampling theory with other factor theories he observes that "if we sampled the *whole* pool of a mind, we should again find the tendency to hierarchical order. If the mind is *organized* into sub-pools (such as the verbal sub-pool, say), then we shall be liable to fish in two or three of them, and get a rank of 2 or 3 in our matrix, i.e. get two or three common factors, in the language of the other theory" (9, p. 279). What Professor Thomson calls "sub-pools" in his sampling theory may possibly correspond to what I have called primary factors. As regards the possibility of innate factors, he states that Spearman considers the general factor g to be an index of the span of the whole mind. Thomson believes that other common factors measure only sub-pools, linkages among bonds. "The former measures the whole number of bonds; the latter indicate the degree of structure among them. Some of this 'structure' is no doubt innate; but more of it is probably due to environment and education and life" (9, p. 283).

These later quotations would seem to indicate that Professor Thomson admits the possibility of some degree of structure in the form of linkages or sub-pools in his sampling theory, that these linked groupings of elements may appear as primary factors in a simple structure such as the verbal factor, and that such factors can be innate or acquired, or both. Even though Professor Thomson makes these particular statements, his main discussion is clearly

opposed to the psychological interpretation of factors, and he repeatedly expresses doubt about the possibility of psychological uniqueness of factors. He does not, however, deny this possibility, for in the chapter which he devotes to the concept of simple structure he says: ". . . and in any case experiment has not yet produced many batteries which clearly exhibit this phenomenon. Undoubtedly, however, if more such batteries were produced . . ., and if the resulting 'simple structure' factors were compatible with one another and in fair agreement with psychological intuition, there would be formed an apparatus for defining factors which would have considerable influence on the progress of psychology" (9, p. 251).

Simple structure actually has been found for different test batteries on different populations, and the primary factors from different experiments actually are compatible with one another. The results are not perfect, but the findings already available make the interpretation of primary factors one of the most challenging tasks in psychology.

As far as I understand Professor Thomson's sampling theory, it can be extended to cover the primary factors in a simple structure by conceiving the universal elements in that theory to be more or less structured into groups, or what he calls "sub-pools." I have not attempted any such statistical conception about the underlying character of the primary factors because I have the suspicion that the primaries may differ radically among themselves in their fundamental nature. Some of the primary factors may turn out to be defined by endocrinological effects; others may be defined in biochemical or biophysical parameters of the body fluids or of the central nervous system; other primaries may be defined by neurological or vascular relations in some anatomical locus; still others may involve parameters in the dynamics of the autonomic system; still others may be defined in terms of experience and schooling. I get more confusion than gain by superimposing an abstract statistical system of universal mind-elements upon the search for the underlying nature of the primary factors. I seem to have little interest in postulating, in a statistical and abstract manner, an undifferentiated and uniform field of mind-elements or neurological machines which combine their activities in summation to produce test scores. That sort of preconception is probably not necessary, and it may not be helpful. I prefer to locate a functional unity in a simple structure without making any statistical hypotheses about its character, to determine whether the functional unity reappears in successive experiments, to make some

guess about the kind of task, or behavior, or end-product which seems to involve the factor, to name it tentatively in terms of its observable end-product, to guess separately for each factor about the underlying processes which identify it, and to head as soon as possible to direct forms of laboratory experiments in terms of which the primary factors may eventually be better understood. When they are eventually identified in some fundamental way, they will no longer be known as factors. They will then be absorbed in the concepts that govern the functions involved. The study of several primary factors may even involve quite different laboratory methods or even different biological and physical sciences. It would probably be a handicap, in running down the fundamental nature of several totally different primary factors, if we had to drag along some preconception about statistical mind-elements or mind-engines in the abstract. The method of factor analysis implies nothing about the biological, or physical, or statistical character of the primary factors.

It is clear, however, that my interpretation of the factor problem is determined in large part by my own preconception, namely, that mentality is not an undifferentiated mass, but that it functions in terms of differentiable processes which do not all participate with equal prominence in everything that mind does.

Invariance of the Factors in a Test

In several previous publications I have stated what I consider to be one of the important principles in factor analysis for the discovery of primary factors. This principle is that *the factorial composition of a test must remain invariant when the test is moved from one battery to another which involves the same common factors* (**12**, p. 120). The reason why I discarded some of my earlier solutions to the factor problem was that they did not satisfy this fundamental requirement. I found one solution that satisfied this requirement in the rotation of the reference frame to the position which shows what I have called a simple structure or configuration. Other solutions can perhaps be found, but I have not seen any other solution that satisfies this requirement.

Since any factor method must satisfy this requirement in order to be acceptable, it is perhaps worth while to elaborate this principle further to show what is implied and what is not implied. It should be clear, first, that the principle applies to different test batteries drawn from a larger pool of tests that have been given to the same population. In such a situation any particular test which is moved

from one battery to another for successive factoring must retain its factorial composition, independently of the battery with which it is associated, as long as each battery contains the common factors that characterize the test. If the test is moved into a test battery which lacks one or more of the common factors that characterize the test, then, of course, the corresponding part of the common-factor variance of the test moves into the unique variance of the test and is lost in other specific factors. Hence, the invariance of the factorial composition of the test demands that the successive test batteries into which the test is inserted must have the same common factors.

In practice, it should also be found experimentally that the factorial composition of a test remains practically invariant when it is inserted in several test batteries that have been given to different, but comparable, populations. The comparability of populations is sometimes a debatable matter, but if the several populations are drawn as samples from the same universe, the principle should apply, though not so accurately as when the several test batteries are all drawn from a pool of tests that have been given to the same population. If a test is given to several populations of different ages or education, then the factorial composition may be expected to change.

General Factors

In several factor studies on adults we have not found any general factor that might correspond to Spearman's " g," but one factor study of 10-year-old children has revealed a general factor in addition to the primary factors that have been found for adults (22). This raises the interesting question whether a general factor will be found among children of other ages and whether it tends to disappear in adults. A large study of 60 tests on 1100 eighth-grade children has just been analyzed. It does not reveal any general factor in the factor matrix F, but a second-order general factor is indicated in the correlations of the primary factors.

If a general factor should be found consistently with young children and not with adults, some attempt must be made to rationalize the finding. One of the possible explanations that come to mind is that the general factor, under these circumstances, might be a maturation effect that disappears when the experimental population is composed of adults. There might be individual differences in the rate at which mentality develops, and this rate might be more or less independent of the adult level of the abilities. Two children might differ because of rate of development even though they will be com-

parable in abilities as adults. To the extent that the rate of development is more or less independent of adult level, the intelligence quotients will show a drift up or down in value with increasing age toward maturity. In the present example, let us assume the simple case in which all of a child's abilities mature at approximately the same age. This is an approximation that can be assumed here to illustrate the manner in which maturation might appear as a general factor.

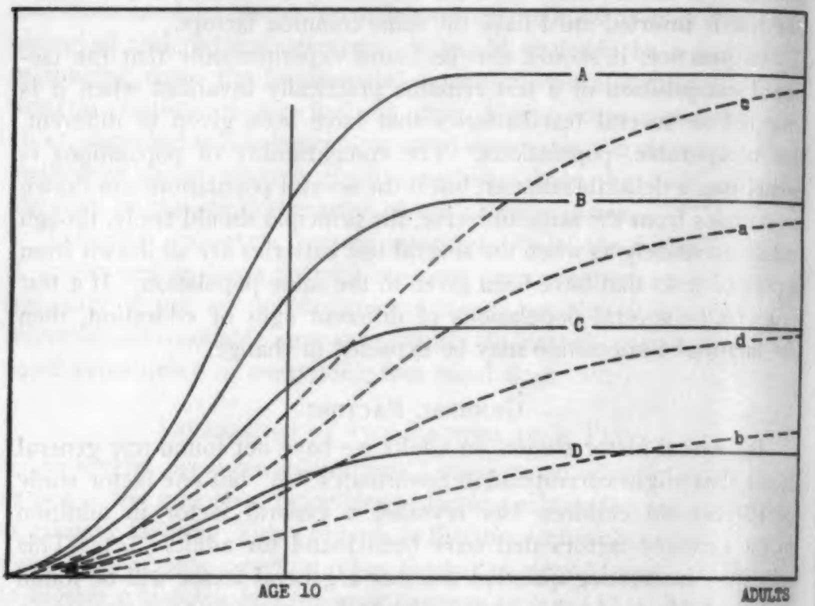

FIGURE 5

Let the solid lines in Figure 5 represent the development of four primary abilities of one child, and let his adult levels in these primaries be denoted by the capitals *A, B, C, D* in the figure. All of these four solid lines are drawn so as to approach their respective adult levels at the same rate. The dotted lines represent the development of another child in the same four abilities, and the rate of development is assumed to be slower. The two children approach the same average adult level. The slower child's adult performances are represented by lower-case letters *a, b, c, d*. The adult levels are drawn so as to represent zero correlation between the primary abilities. Now if the appraisal is made at an earlier age, say 10, it is clear from the figure that there will be strong correlation among the primary factors because

of a general maturation factor. The child whose mental development is slower will have relatively low scores in all of the tests, whereas the child of faster development will have higher scores in all of his tests. The result will be a common factor that is due to individual differences in rate of maturation. This factor may dominate the variances of the tests. In addition to the general maturation factor, there should be the same primary factors as for adults. If the same group of children were tested again 10 or 15 years later, the maturation factor should have disappeared. This hypothesis is offered only as one of the possible explanations if it should continue to be found that a general factor appears in data for children and not for adults.

The concept of simple structure is misinterpreted in curious ways. In a recent monograph (**16**) I made the statement that, so far, we have not found the general factor of Spearman, but that our methods do not preclude it. In discussing the concept of simple structure, Zubin (**23**) says: "One necessary result of the rotation of axes and the maximizing of the number of zero factor loadings is the elimination of the general factor running through all the tests which Spearman postulates. . . If conscious rotation to eliminate the general factor does not preclude it, what will?" This is a hasty conclusion. As a matter of fact, there are at least two different kinds of general factors that can be revealed by the methods which I have been using. If the primary factors are all related by means of a common general factor, then the primary factors are correlated. The matrix of their intercorrelations is then of rank 1, and the general factor can be determined directly. If there are n such correlated primary factors, the new resolution will have $(n+1)$ orthogonal factors, namely: one common general factor and n orthogonal primary factors. If the general factor is present in all of the tests, but is uncorrelated with the primary factors, then the primary factors can be found, as in the first case; but there will be a large residual variance which is unaccounted for by the primary factors that determine bounding planes of the configuration. It seems repetitive to say again that there is no guarantee that any given correlation matrix will show a simple structure and that there is no guarantee that the primary factors which determine bounding planes will account for all of the variance of the tests. If the remaining variance can be accounted for by a factor which is orthogonal to the primary factors that have been isolated by the bounding planes, then such a factor may be found to be present in all of the tests, and it may be a functionally significant general factor. But it may also be a composite of several

primary factors that are not clearly represented in the tests of the battery. Only further experimental investigation can answer such a question. The first type of general factor can be determined with more certainty in a single experiment than the second type of general factor.

We must guard against the simple, but common, error of merely taking a first centroid factor, a first principal component, or other mean factor, in a test battery and then calling it a general factor. Such a factor can always be found routinely for any set of positively correlated tests, and it means nothing more or less than the average of all the abilities called for by the battery as a whole. Consequently, it varies from one battery to another and has no fundamental psychological significance beyond the arbitrary collection of tests that anyone happens to put together for a factor analysis.

I have just read Professor Spearman's latest paper (3). In that paper he does exactly what I have described here. He takes the average of all the abilities in the 56 tests that I happened to assemble for a factor analysis, and he calls this average factor "g." It has a saturation in each of the tests. By assembling another battery with many more tests of one kind and fewer tests of another kind, we could alter the saturation of such a " g " factor within wide limits. The saturation of any particular test with that kind of " g " can be raised or lowered by merely putting the test in different arbitrary test batteries. This elementary point in factor analysis has been made repeatedly in previous publications. The fact that the average of all factors in a test battery can easily be found even by a novice in factor analysis does not show that the average factor has any psychological significance beyond the arbitrary collection of tests for which such a factor is just an ordinary average. As psychologists we cannot be interested in a general factor which is only the average of any random collection of tests. One or more general factors may exist in the sense of psychological invariants, and we are still attempting to find them.

Factors and Faculty Psychology

One of the common criticisms of the psychological interpretation of primary factors is that such attempts are faculty psychology. I have insisted that factor analysis must make sense. I am not satisfied with mere statistical descriptions of correlations and test scores. I am not satisfied with factorial condensations of data unless they help us to solve some psychological problem. What do we learn for

psychology in a factor analysis? For what psychological problem was the test battery devised? Was the factor method, with its assumptions, adequate for that particular psychological problem? Except in methodological studies, these are the first questions to ask in any factor study.

Those who criticize factor analysis as faculty psychology talk in the next breath about verbal and nonverbal intelligence, special aptitudes for music and for art, mechanical aptitudes, and disabilities in reading and arithmetic, without realizing that they are implying factor analysis and the interpretation of factors. Conversationally, we can talk about musical ability in the singular, but, as psychologists, we believe that different abilities are probably involved in a good voice, absolute pitch, originality in harmony, counterpoint, and orchestration, melodic memory, ease in memorizing at the piano, ability in musical interpretation, and so on. How many important functional unities might there be in this domain? They are surely not all completely independent. They are correlated, and they probably represent a smaller number of abilities than the infinite number of musical tasks that could be given to a music student. To find these functional unities is the problem of factor analysis. If that is faculty psychology, then so is most investigation of individual differences and most of the work on special aptitudes and defects. And so are also many of the chapter headings in the textbooks of psychology which the same critics are teaching. The superficiality of the verbalism about faculty psychology does not deserve more than passing notice.

Appraisal of Individual Subjects

So far we have considered the psychological interpretation of the primary factors. Another aspect of factor analysis that is frequently misinterpreted is the description of the subjects in terms of the factors. Suppose that a test shows appreciable variance on two factors, A and B. We cannot infer that all subjects in the experimental population used both of the abilities, A and B, to do the task. Some of them may have used mostly A; others may have used mostly B; still others may have used both A and B. The result will be a factorial composition of both A and B in the test which will have a complexity of 2. How, then, can we ever infer anything about the abilities of the subjects? In exactly the same way as in any other test situation where the test is not a perfect measure of what we really want to measure. Let us use, say, three tests to appraise factor A. Let the tests have factorial compositions AB, ACD, and $AEFK$, respectively. The combination

of the three tests with a factor A in common will be the means of appraising the ability A. This is just what we do in all psychological and educational tests which have some factors in common with their criterion of validity. The principle is as old as mental and educational measurement. To the extent that psychological tests can be reduced in complexity we shall be able to make individual appraisals with more certainty.

If it is known that a task can be done by two different methods which call for different abilities, A and B, then the test should be revised, if possible, so as to make the task less ambiguous as to the method of doing it. Individual testing has the advantage in that the examiner can ask the subject to explain how he solves the test problems. It might be possible to make separate factor analyses of some groups of tasks in terms of the methods voluntarily adopted by the subjects. These procedures would be feasible but laborious. The revision of tests is sometimes indicated by discussing with a number of subjects how they do the tasks.

It seems rather obvious, but it is not generally recognized, that the factorial composition of a test depends on what goes on in the minds of the subjects. The printed test blank, or the laboratory task, is not a psychological entity with a fixed factorial composition for all people of all ages on all occasions. While number problems may be routine tasks for an adult, they may be inductive tasks for a child. Or the factorial composition of a test may be altered for the same subject if he is tested before and after some particular form of instruction. In a review of Anastasi's attempts to use the factor methods I illustrated this possibility by a syllogism test (17). As ordinarily given, it may have practically no part of its variance on the space factor. But if you teach the subjects how to use Euler's circles for solving syllogistic problems and then try the same test again, the retest may show an appreciable variance in the space factor. Many of these difficulties would disappear if the critics would think of factor analysis as an integral part of a psychological problem rather than as an isolated and self-contained statistical routine that grinds a psychological problem into mathematical artifacts. Even a good tool will ruin things if it cuts blindly without regard to the nature of the job that is to be done.

The development of tests for the primary mental abilities naturally raises the question of their application to vocational guidance. The improvement of such a test consists in reducing its complexity so that a large part of its variance is attributable to only one primary

factor. But this purification of a test reduces its correlation with a complex function such as professional success. This is, in one sense, the opposite of the vocational aptitude tests which are intended to give high correlations with professional success. If an occupation depends on, say, 10 factors, the purification of a test so that it confines its appraisal to only one of the 10 factors will probably reduce its correlation with occupational success. The solution would be to use separate tests for the 10 primary factors and to use a linear combination of them for the prediction of occupational fitness. In some cases it will be possible to appraise a particular occupational aptitude more readily by a short composite and complex test which is assembled specially for a particular occupation. In general, it will be more profitable to use the purified tests for the primary factors, because the resulting profile should enable us to make predictions for a wide range of occupations that involve the primary factors in various combinations. This is another aspect of the difference between the crude empirical approach to the vocational guidance problem and a rational approach to it.

Terminology

Professor Tryon objects strenuously to factor analysis, and especially to the " orthodox " factor analysts, as he calls some of us. He says, in a recent lithoprinted publication (**18**, p. 2): " The other way is called *orthometric analysis,* which is a new term for the method called *factor analysis,* this latter being a bad term as it implies that a mere statistical technique proves the existence of *factors* in behavior." He objects not only to the term " factor," but he also ridicules all attempts to interpret factors psychologically. Quoting him further (**18**, p. 2): " The search for general components is a popular pursuit. There have appeared in the psychological literature thousands of tables of correlations between intellectual and personality measures, in the analysis of which investigators have sought to give a coherent statement of so-called *underlying* general factors or radicals at work in the various measures. The results have been confusing and the conclusions unconvincing, largely because the statistical techniques of analysis employed have had imputed to them powers they do not possess, namely those of isolating 'psychological factors.' " " The psychologist is asked to discover radical or primary components. But unhappily a study of the intercorrelations between measures cannot reveal these. He can, nevertheless, discover general components which act *as if* they were unitary determiners in different behaviors.

Such determiners are here called *operational unities.* These are defined as those components which result in two or more variables showing the same pattern of correlation coefficients with all the other variables in an investigation. Two variables, *A* and *B*, are said to be wholly or partially determined by an operational unity if both correlate high with variable *M*, low with *N*, intermediate with *O*, and so on throughout the other variables. In such a case, clearly what is general in *A* and *B* is behaving in an identical and unitary fashion." These operational unities of Tryon sound suspiciously like the primary factors, and one wonders whether most of the fire is not about names instead of ideas.

In further elaboration of operational unities, Tryon says (**18**, p. 4): "When an investigator has measured a group of *N* subjects by means of *n* psychological measures and has calculated all of the intercorrelations between these variables, his first problem is that of finding how many operational unities appear to exist." But that is the ordinary factor problem of determining the number of common factors in a test battery (**12**, p. 72). Again he says (**18**, p. 6): "If we wish to name an operational unity, we proceed by noting what appears to be the psychological content common to the variables in a cluster and in other variables with which they show high correlation, and at the same time absent from those variables with which they show low correlations." "The names simply attempt to describe what appears to be operating in a similar and unitary fashion in the variables of a cluster group." But this is just the procedure in naming the primary factors, which Tryon ridicules (**16**, Chap. V). And he says further (**18**, p. 8): "The discovery of an operational unity is, however, of considerable importance on a number of counts. It contributes some unity to the data which might not have been suspected. It provokes important psychological speculation as to possible radical components at work, whose operation might not otherwise have been discerned." But this also is the principal purpose in finding primary factors whose underlying processes should be determined. He endorses the "discovery" of operational unities, but denounces the same process for primary factors. He says (**18**, p. 30): "The writer believes that the so-called 'verbal,' 'space,' 'number,' etc., factors so frequently 'discovered' by factor analysts are such operational unities." Now, if the verbal, spatial, and number factors are operational unities, and if it is worth while to find operational unities, then our principal error, according to Tryon, seems to be that we have used the wrong term, primary factor, instead of

operational unity. I must have come pretty close to the right term when I referred to primary factors as functional unities.

The elemental components, which Tryon conceives to be the determiners of human traits, he calls radical components (**18**, p. 1). "Radical components are causal, primary, univocal, indivisible determiners of differences. It is certainly conceivable that radicals exist." Tryon does not attempt to isolate these hypothetical elemental components, and I do not believe that any student of factor analysis has seriously attempted to discover any of these ultimate indivisible elements which might be conceived to produce individual differences in human traits.

In one of my earlier papers I used the term "unitary trait" (**11**). To some readers this term might mean a trait that is either completely present or entirely absent in an all-or-none sense. The term can also mean a grouping of correlated effects which act together as a single variable trait. Such a trait shows a continuous distribution in the population, while an elemental all-or-none trait (an attribute) shows a two-point distribution. The Committee on Unitary Traits of Professor Spearman used the term for continuously variable traits. This is the way in which I used the term formerly. This seems also to be the sense in which Professor Tryon uses the term. During the last five years I have differentiated these two kinds of traits by the term "primary" for basic, continuously variable traits and the term "unitary" for the hypothetical, elemental, all-or-none attributes which Professor Tryon has called radicals. The elemental individual attributes are of only speculative interest because no factorial studies have as yet revealed any such elemental attributes in psychology.

Clusters

When we inspect a table of intercorrelations in search of a small number of basic categories, the first natural idea is to see if the test variables group themselves. One hopes, then, that several of the tests, say verbal tests, may be highly correlated among themselves and that they agree in the list of tests with which they have low correlations. One can usually find such clusters of like-behaving variables in a correlation table. My first attempts in factor analysis were of this kind, and I devised several routine procedures for determining clusters (**12**, p. 174). The conspicuous clusters were easily determined in several ways, but the clusters were not unique. They have a way of sliding over and through each other that is annoying. Professor Tryon has recently proposed that we return to cluster analysis

as a principal factor method. It is not likely that serious students of the factor problem will accept Professor Tryon's suggestion except for rough inspectional purposes.

One of the principal difficulties with cluster analysis can be illustrated by a simple case. Suppose that two or three parallel forms of each test are included in the test battery. Then Professor Tryon's procedure would show as many "operational unities" or clusters as there are tests, because surely the parallel forms would be like-behaving in the correlation matrix. By such methods any test whatever can be an operational unity. Such a situation is not likely to be very revealing psychologically. In my own work such doublet factors are discarded routinely. If discrimination is to be made by the investigator's judgment as to when several tests are parallel or distinct, then the objectivity—upon which Professor Tryon is so insistent—is lost.

Communalities

There is a difference of opinion among students of factor analysis as to what values should be recorded in the diagonal cells of the correlation matrix. Some investigators write unity in the diagonal cells. In studying this problem I came to the conclusion that what I have called the common-factor variance should be recorded in the diagonal cells of the correlation matrix, and I coined the name "communality" for these values (**10**, p. 8; **12**, p. 62). They are somewhat lower than the test reliabilities. It is computationally awkward that the communalities are unknown at the start of an analysis; hence, they must be estimated. The best estimate is obtained by adjusting the communalities for each factor that is extracted from the correlation matrix. Professor Kelley and Professor Hotelling do not accept the communalities as the correct diagonal entries. They use unity or the reliabilities in these cells. I have elsewhere discussed this problem in some detail, and I doubt whether I could restate the issue any more clearly here (**12**, Chap. I). Some investigators who have worked for meaningful factors use the communalities, and Professor E. B. Wilson seems to take them for granted in a recent mathematical paper (**19**).

In stating his objections to the use of communalities Professor Kelley considers a hypothetical battery of 10 tests, two of which involve a certain music capacity not present in the other eight tests. The communalities of the music tests include their variance in the music factor, since it is common to two tests in the battery. If, now,

one of the music tests be dropped from the battery, the other music test will have a reduced communality. The decrement is due to the music factor, which then becomes a specific factor and no longer a common factor. "Capacity in music reflected in part in the one test retained is now specificity, and according to Thurstone not worthy of retention in the analytical picture" (2, p. 56). It is, of course, not a question whether the specific music ability is "worthy" of retention. The question is how to disentangle the music factor from other specific factors in the same test and from the error variance in that test. The separation of the specific music factor from the other specific factors in the remaining musical test is not solved by Professor Kelley's factorial methods. But the problem can be solved by using an augmented test battery and by rotation of axes in the common-factor space.

Professor Tryon describes three methods of estimating the communalities. First, he describes the procedure of estimating the communality as the highest correlation in the column of the correlation matrix. Second, he describes the procedure of estimating the communality by the intercorrelations of a cluster. And then (**18**, p. 11): "Still another method, which we do not use because it is the hard way, is that of Thurstone. His method requires the laborious operations of his *centroid* analysis . . ." In *The vectors of mind* I described eight methods of estimating the communalities. Method No. 4 (p. 89) is called "Highest coefficient in each column," method No. 2 is called "Grouping of similar tests," and method No. 3 is called "Grouping of three tests." Professor Tryon has adopted these methods, and now he contrasts them with what he calls "Thurstone's Method." Furthermore, it just so happens that the method which he has adopted, the highest coefficient in each column, is the very method that is used in the centroid analysis (**12**, p. 108)!

Factor Analysis and Multiple Correlation

Although there are some interesting mathematical relations between multiple correlation theory and factor analysis, we must not lose sight of the fact that the purposes of the two types of analysis are radically different. If we have a set of measures for each member of a population, and if one of these measures is to be predicted from others in the same population or in comparable populations, then the usual regression equation is called for. It gives the best possible prediction of the dependent variable within the assumptions of linearity. The correlation between the pool of independent variables

and the dependent variable which is to be predicted is, in fact, the multiple correlation coefficient.

The purpose of factor analysis is quite different. If we have a set of measures for each member of a population, and if we want to know whether these measures are related by some underlying order which will simplify our comprehension of the whole set of measures, then a factor analysis is called for. The question whether a factor analysis is to be made in any particular study can be answered by ascertaining whether the purpose is to predict one of the measures by the rest. If so, then the regression equation is indicated, and we can forget about factor analysis. The purpose of a factor analysis is to discover basic and underlying functional unities. These may eventually be used for predicting a wide range of things, but that is not the immediate object of any particular factorial study. The object is to discover basic categories in terms of which we can think of the complex measures that are obtainable experimentally. The factor methods are not immediately concerned with predicting or duplicating either the test scores or the correlations. The correlations are used merely as a guide in discovering what the underlying functional unities may be.

It may be a surprise for some students to learn that factor analysis is not immediately concerned with the prediction or duplication of test scores and correlations. A factor analysis might reveal the existence of a few primary factors which represent only a part of the common-factor variance of the measures. The measures may involve so many factors that none of them can be predicted or appraised individually by the test battery. But such a result might be far more important and fundamental than a method of predicting some test scores from the other test scores. The latter is a trivial matter in comparison with the isolation of functional groupings in the abilities. To identify a primary factor is an important finding even if none of the separate tests can be used for reliable individual appraisal of that factor. But if a primary factor can be identified, then it can eventually be appraised individually with improved measures.

The distinction between factor analysis and the regression equation for the prediction of something is involved in Professor Godfrey Thomson's arguments concerning vocational guidance. If a set of tests has been found to be predictive of success in a particular occupation, then the tests might as well be used directly without translating them into factors. I quite agree with Professor Thomson's conclusion

that "when vocational guidance proceeds by giving to a candidate a number of tests which have previously been given to persons already engaged in the occupation, the use of factors has no mathematical justification whatever" (**6**, p. 53). If factor analysis, or any other method, has revealed the nature of some of the fundamental abilities which are involved in an occupation, then the tests for those abilities will enable us to make predictions far beyond the particular occupation for which a particular test battery has been constructed. As I have stated, the difference is simply that of a crude empirical procedure and a rational one. If we have no landmarks for human abilities, then we can proceed by trying empirically all sorts of tests for all sorts of occupations. Here and there we will hit it when a test battery does serve to predict success in a particular occupation. The rational approach to the problems of vocational guidance is to search for the fundamental abilities and traits which determine various kinds of success. If we knew the nature of these traits, then we could observe the occupations in terms of these traits, and the test batteries could be set up rationally with greater expectation of satisfactory predictions. In testing the predictive value of the tests we should use the ordinary regression equation. In looking for the basic traits we should use the factor methods or other equivalent methods. But Professor Thomson probably would not allow this possibility since he does not believe that human abilities and traits act in unique constellations that can be identified. It seems difficult to conceive just how individual differences in particular aptitudes can exist at all unless mind has some kind of structure that can be described by a system of independent parameters.

In one of his papers concerning the vocational guidance argument Professor Thomson reviews my treatment of the unique variance of each test (**7**, p. 258). In developing the factor theorems I divided the unique variance into two possible sources, the specific factors and the error variance. Professor Thomson is correct in calling attention to the fact that the sampling errors are not unique for each test. Mr. Ledyard Tucker has suggested that instead of dividing the general factor matrix into three sections, as I have done, for the common factors, the specific factors, and the error variance, it would be clearer and more correct to introduce a fourth section to represent the correlated parts of the sampling errors and the variable errors. In general, this section would be filled. It would not be a diagonal matrix. The common factors are oblique to the extent that sampling

errors and chance errors in the scores have introduced slight correlations between the primary factors which might be orthogonal in the general population. Fortunately, the only difference is in the exposition of the subject. The computational procedures are in no way affected, nor is there any effect on the interpretation of the resulting oblique primary factors. Mr. Tucker will present this revised development in a separate paper.

Methods of Factoring the Correlation Matrix

A factor analysis for the isolation of primary mental abilities has, generally, two separate stages, namely: first, the factoring of the correlation matrix, and, second, the rotation of the reference frame in the common-factor space to determine whether a simple configuration exists in the test battery. The purpose of the first stage is to separate the common factors from the specific factors and error variance. If computational labor were not a consideration in this work, the ideal procedure, as far as we know the problem at the present time, would be to extract the main principal axis with estimated communalities in the diagonal cells. The residual table would be treated in the same way, namely, by extracting the main principal axis from the residual table with adjusted diagonal entries. By this procedure the common factors would be separated from the specific and unique factors as well as it could be done now. In this manner the unique variance which is represented in the diagonal entries would not be allowed to dominate the factor matrix, as is the case when the principal axes are determined without diagonal adjustment for the communalities. The first solution that I devised was the principal axes solution without diagonal adjustment for each successive factor, but I discarded it because it did not separate satisfactorily the common-factor variance of the tests (**10**, p. 17; **12**, Chap. IV).

The next solution that I devised was the centroid method, which is a compromise with what I should consider to be the ideal solution (**12**, Chap. III). It requires considerably less labor, but I shall gladly discard it when a more efficient method is devised. The centroid method has served well and gives a good separation of the common factors from the unique factors.

There is a possibility that these computations can be very much accelerated by means of punched cards and tabulating equipment. Several such procedures are now being tried. It is also possible that the plotting of the diagrams can be accomplished on the tabulating machines. In addition, several other mechanical devices are being

tried. Mr. Ledyard Tucker has designed an ingenious and simple mechanical device for making the sign changes in the centroid method. A set of transformers is now being built which may solve the computational problems that involve multiplication of large matrices.[2] Professor Woodrow has suggested a method of determining a simple configuration directly from the correlation matrix (**21**).

Elsewhere I have shown that it is futile to attempt the psychological interpretation of the centroid matrix without rotation or of the principal axis matrix without rotation (**12**, Chap. VI; **13**). This is not now so often attempted as a few years ago.

In one of my large factorial studies I used tetrachoric correlations for the correlation matrix to save computational labor (**16**). For this I have been criticized. If both variables in a correlation table are normal, it is not likely that the correlation surface will depart much from normality. It is not a violent assumption to deal with the surface as if it were a normal surface. However, I should have made explicit the assumption of a normal correlation surface in the study in which tetrachoric coefficients were used. The greater instability of the tetrachoric coefficients does not seriously affect a factor problem in which a large number of variables is involved. The criticism against tetrachoric coefficients is, however, easily met by using product-moment coefficients instead. We have used product-moment coefficients in all of our test studies except the first one.

In applying statistical concepts to these problems we must not forget the nature of the assumption on which the statistical methods rest. Factor analysis starts with the assumption that the score in a test can be regarded as a linear combination of hypothetical factor scores. The whole theory is built on this assumption. But it would be just about as good to assume that the square of each score is equal to a linear combination of factor scores, or the logarithm of the score. Almost any monotonic function of the scores could be set equal to the linear combination of factor scores. The way in which a test is scored is, after all, more or less arbitrary; and the subject's score depends also on the distribution of difficulty of the items, whether they are arranged in increasing order of difficulty, and so on. Fortunately, the factor methods can make sense out of a great variety of

[2] Since this manuscript was written, we have transferred successfully the computations for the centroid analysis to punched-card methods, and the diagrams are now plotted on the tabulating machines. A modified form of scoring machine has been constructed for matrix multiplication with a capacity of 15 columns. These new methods will be published soon.

such assumptions, and that is why we can find the primary factors involved in a test battery. If a simple structure exists in a battery, one would probably find it also if the whole job were done over again from the start with the logarithms of the scores instead of the scores as they happened to be recorded. The correlations would be different, but the simple structure would probably be revealed as well in one case as in the other. Hair-splitting is likely to be misplaced if it follows upon coarse initial assumptions.

Minimizing the Number of Common Factors

Professor Godfrey Thomson has discussed in several publications the maximizing of the specific factors which he describes as a central principle in Spearman's factorial methods and in mine (7; **8**, Chap. VIII). Professor Spearman, in his principal work, is interested in one central and universal factor; and he naturally determines it, in any particular test battery, so that the central and universal factor which he postulates will account for as much as possible of the test variances. One might say, then, that Spearman *minimizes* the specific factors, subject to the condition that there shall be one commanding universal factor in the battery. Whether it is maximizing or minimizing depends somewhat on how one puts the problem.

In discussing my multiple factor analyses Professor Thomson starts by noting that I try to account for the test correlations with the *minimum* number of common factors, and such a statement leads to the conclusion that I *maximize* the specific factors. It is true that I have described the factor problem by saying that we want to account for the intercorrelations by the smallest number of common factors. In a careful exposition one should foresee, if possible, the different ways in which the exposition can be misunderstood, and this is a case in which I have failed to make this matter clear.

The variance of each test has two complementary parts, the common-factor variance that is common to two or more tests, and the residual singular variance in each test. Fortunately for the factor problem, the singular variance is confined to the diagonal cells, and the common variance is represented in the side entries. We start with a correlation matrix in which the diagonal entries are unknown. Now, *the number of common factors is the rank of the matrix as determined by the side entries.* Strictly speaking, it is not a question of maximizing or minimizing anything. The matrix, as it stands, has a certain rank, determined by the side entries. That rank determines also the diagonal entries which then become the communalities.

If any higher rank is to be imposed, the increment in rank derives from the diagonal entries because the side entries are experimentally given. A lower rank is impossible because of the side entries. The problem can also be put in another way. Let the diagonal entries be inserted. If they are too low, the matrix will not be positive definite. Hence, it will not be a correlation matrix. If they are too high, the rank will be higher than that which is determined solely by the side entries. The communalities in the diagonals have values such that lower values destroy the Gramian properties of the matrix and higher values increase the rank beyond that of the side entries. In order to find this rank, which is determined by the intercorrelational side entries (not the self-correlations), we can say that the number of factors is the minimum rank that can be found for any set of positive diagonal entries between zero and unity which do not destroy the Gramian properties of the matrix. While this states the problem operationally, it is not because we are trying to minimize anything. The rank is there in the side entries, and it is actually the number of common factors that functioned in producing the correlations.

Professor Thomson's description is that " the procedure has made the number of common factors as small as possible . . . *but at the expense of making the n specifics as large as possible*" (7, p. 256). The procedure is, rather, to determine what the number of common factors *actually is* in the correlation matrix. It is fixed for us already in the correlation matrix. We cannot maximize it or minimize it if the problem is determinate (12, p. 76).

If it happens that the unique test variances are large, then it simply shows that a large part of the tests remains still unaccounted for. No statistical jugglery in maximizing something will help here. We must try again with new batteries to change the unique variances to common-factor variances until the unique variances reduce to measurement and sampling errors. Each correlation matrix is, of course, analyzed independently. If the new test batteries are successful, then the communalities will rise, not because of anything that we maximize in statistical manipulation, but because the common-factor variances have actually increased, leaving a smaller part of the variances unknown. Such improvement and reduction of the unknown specifics is not accomplished by any statistical or factorial procedure. It is accomplished by the investigator if he has good psychological intuitions in setting up his experiments.

Factorial Study of Psychological Hypotheses

In discussing the interpretation of psychological factors we have considered a few examples of the manner in which the factor methods can be used to investigate psychological hypotheses. In the first investigations of any domain by the factor methods one proceeds with only tentative classifications of a large number of measures in the field, in the hope of finding at least a few dimensions that show a clear simple structure. It is likely that the first study of any domain will show only a few clear primaries and that a large part of the common-factor variance and specific variance will remain obscure as regards the configuration of the test battery. When the clearly indicated primary factors have been given tentative interpretation, they should be represented in new test batteries, and new tests should be devised which represent the tentative factors as strongly as possible. The new test batteries should be factored in order to determine whether the new tests show large saturations in the tentatively identified factors. If they do not show this result, then the interpretation of the primary factors is not sustained, and a new hypothesis must be devised as to the fundamental nature of the processes that characterize the new primary factors. If the new tests do show large saturation in the expected primary factors, then the hypothesis concerning the nature of the primary is sustained. But factorial experiments are no exception to the general rules for all scientific experimentation. One can demonstrate that an hypothesis is wrong if it does not agree with experimental results, but one can never demonstrate that an hypothesis is correct. Some new experiment might be devised to show that the former experiment covered only some special case. The interpretation of factorial results is subject to the same limitations. If a primary factor is verified in repeated experiments, it is always possible that some future experiment will reveal that a primary factor consists really of two or three primary factors which happened to work together so that they were not separated in the earlier experiments.

Correlated Primary Factors

Among statisticians and psychologists there is a rather general belief that if human traits are to be accounted for by any kind of factors, then these factors must be uncorrelated. The geometrical representation of uncorrelated factors is a set of orthogonal reference vectors. This belief has its origin in the statistical and mathematical convenience of uncorrelated factors and also in our ignorance of the nature of the underlying structure of mental traits. Since we know so

little about them, and since it is statistically convenient to use uncorrelated reference traits, the insistence on orthogonality can be understood, but it cannot be justified. Height and weight are two useful measures of body size even though they are correlated. If we should insist on using instead two linear combinations of height and weight which are uncorrelated in the general population, we might find such measures statistically more convenient in some situations, but they might be awkward to think about, and there might not be any unique set of orthogonal measures to cover these traits. We might as well

TABLE I

	Dimensions		
Boxes	x	y	z
1	3	2	1
2	3	2	2
3	3	3	1
4	3	3	2
5	3	3	3
6	4	2	1
7	4	2	2
8	4	3	1
9	4	3	2
10	4	3	3
11	4	4	1
12	4	4	2
13	4	4	3
14	5	2	1
15	5	2	2
16	5	3	2
17	5	3	3
18	5	4	1
19	5	4	2
20	5	4	3

make use of the freedom of factor analysis as regards orthogonality by allowing the primary factors to be correlated if they really are correlated.

As a numerical example of correlated primary factors I have chosen a set of relations so well known that no technical knowledge is needed to understand them. A random collection of boxes constitutes the population in this example. Let us imagine a collection of boxes and a set of 20 measurements for each box. Let us assume, as is really the case in most psychological problems, that we have no idea what these measurements represent and that we believe some underlying order may exist in the 20 measurements for each box. In this fictitious example the boxes constitute the individuals in a statistical population, and the measurements correspond to 20 test scores for each individual. In Table I we have the three dimensions,

x, y, z, for each box. The box population is here defined in terms of 20 different rectangular box shapes, and it is assumed that these 20 shapes occur with equal frequency in the population of boxes. For example, if the population consists of 200 boxes, there are 10 of each kind. Any other arrangement could have been taken. In Table II we have a list of the 20 measurements (tests) that were taken for each box. These measurements consist of various nonlinear functions of the three dimensions, x, y, z; but, in the factor analysis, the

TABLE II

Tests	Structure x	Structure y	Structure z	Test Formulae
1	+			x^2
2		+		y^2
3			+	z^2
4	+	+		xy
5	+		+	xz
6		+	+	yz
7	+	+		$\sqrt{x^2+y^2}$
8	+		+	$\sqrt{x^2+z^2}$
9		+	+	$\sqrt{y^2+z^2}$
10	+	+		$2x+2y$
11	+		+	$2x+2z$
12		+	+	$2y+2z$
13	+			$\log x$
14		+		$\log y$
15			+	$\log z$
16	+	+	+	xyz
17	+	+	+	$\sqrt{x^2+y^2+z^2}$
18	+			e^x
19		+		e^y
20			+	e^z

number of basic factors and their combination in the 20 measurements are assumed to be entirely unknown.

It will be seen in the formulae of Table II that each of the tests 1, 2, 3 represents the square of an edge of the box; 4, 5, 6 represent the area of a side; 7, 8, 9 represent the diagonal of a side; 10, 11, 12 represent the perimeter of a side; and the other tests are similar arbitrary functions of the three box measurements. In dealing with the 20 measurements for each box we treat them merely as if they were 20 test scores for each box. We correlate them and factor them as if we knew nothing about the manner in which the scores became correlated. Our object is to discover by factor analysis if the 20 measurements for each box reveal some underlying physical order. Actually, we have set up the measurements so that three factors are

involved—namely, the three dimensions of a box—but we deal with the problem factorially as if we did not even know how many factors were involved in the 20 test scores for each box. In order to simplify the problem, errors of measurement and sampling errors are here omitted. The correlations are, therefore, unaffected by the number of multiples of the 20 different box shapes that constitute the statistical box population. The correlations were determined from the box measurements listed in Table I and the test formulae of Table II. In Table III we have the centroid matrix for three factors. The mean

TABLE III

Centroid Matrix F_c

Tests	I	II	III
1	.659	—.736	.138
2	.725	.180	—.656
3	.665	.537	.500
4	.869	—.209	—.443
5	.834	.182	.508
6	.836	.519	.152
7	.856	—.452	—.269
8	.848	—.426	.320
9	.861	.416	—.299
10	.880	—.341	—.354
11	.889	—.147	.436
12	.875	.485	—.093
13	.667	—.725	.109
14	.717	.246	—.619
15	.634	.501	.522
16	.936	.257	.165
17	.966	—.239	—.083
18	.625	—.720	.166
19	.702	.112	—.650
20	.664	.536	.488

of the absolute residuals (disregarding sign) after three factors had been extracted was .008, which does not warrant the extraction of more factors. At this point we would know that only three factors were involved in the 20 fictitious box scores.

The rotation of the factor matrix is a very simple matter because in the method of extended vectors (15) there is only one plot for three dimensions, and the three primary factors should appear directly in that plot if they exist in the battery. Figure 6 is that plot. The triangular arrangement of the points in the figure is immediately apparent, and the three positive bounding planes were determined from the diagram. The resulting rotated factor matrix is shown in Table IV. A simple structure is apparent in Figure 6 as well as in Table IV.

In the interpretation of the primary factors of Table IV we look for the tests which have the high saturations in the first column. These are tests 1, 13, and 18; and we conclude that the first primary factor is very heavily represented in these tests. In this fictitious problem we can actually turn back to Table II to find the formulae for these tests in terms of the primary factors. These are, of course, not known in psychological problems. We find that tests 1, 13, and 18 are the only three tests which depend exclusively on the x-dimen-

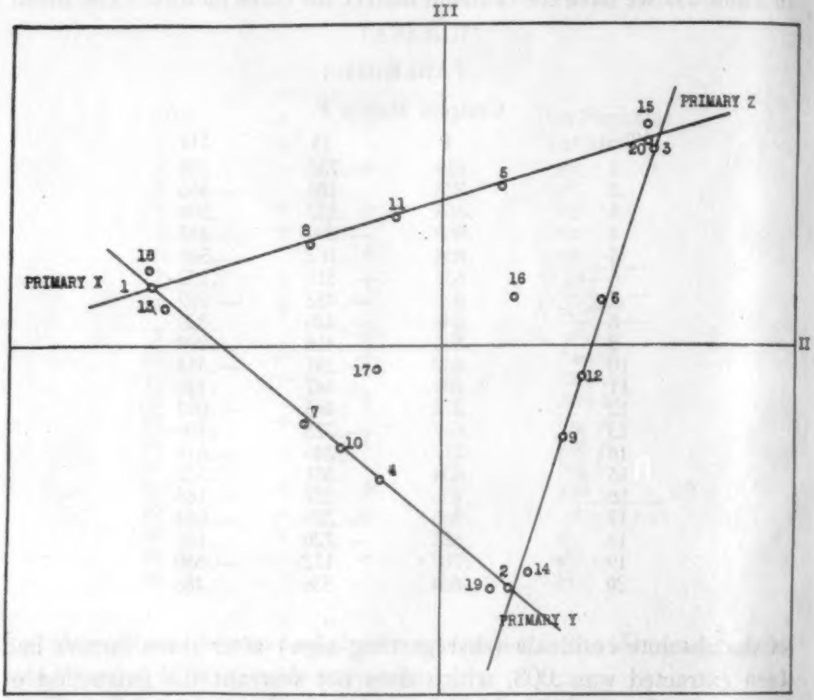

FIGURE 6

sion of the boxes. Further comparison shows that all of the tests which are on the plane YZ in Figure 6 do not contain X in their formulae. The tests in that plane have no part of their variance attributable to the primary factor X.

The same kind of comparison can be made for each of the three primary factors X, Y, and Z. Those tests which depend on only one of the primary factors are in the corresponding corner of the diagram, and they are represented in Table IV with large factor loadings in the proper column. Those tests which are functions of two primary factors are represented as points on the sides of the triangle, and the

points are consistent throughout as to the pairs of primary factors involved. Finally, the two tests 16 and 17, which have formulae containing all three of the primary factors, are inside the triangle. They are the volume and the principal diagonal, respectively. It is interesting that the linear approximations which are assumed in factor analysis reveal a simple configuration, even though the 20 measurements were nonlinear functions of the primaries. The discrepancies are represented by the small entries in Table IV, which should be zero if the factor methods did not depend on the assumption of linearity as a first approximation.

TABLE IV

Rotated Factor Matrix V

Tests	X	Y	Z
1	.969	.003	—.003
2	.025	.939	.005
3	.007	.023	.957
4	.476	.727	.000
5	.387	.005	.844
6	.012	.393	.794
7	.719	.512	—.025
8	.850	.016	.384
9	—.011	.760	.444
10	.615	.623	—.008
11	.668	.008	.638
12	—.007	.610	.628
13	.956	.034	—.013
14	—.024	.921	.063
15	.028	—.019	.937
16	.282	.362	.704
17	.644	.460	.273
18	.947	—.032	.008
19	.072	.906	—.040
20	.004	.032	.948

Next, we note that the three primary factors in this population of boxes are correlated. The correlations among the primary factors as determined by the factor analysis are shown in Table V. The

TABLE V

	Actual Correlations				Obtained Correlations		
	X	Y	Z		X	Y	Z
X	1.00			X	1.000		
Y	.25	1.00		Y	.229	1.000	
Z	.10	.25	1.00	Z	.105	.224	1.000

actual correlations between the primaries in the original data are also shown in the same table, and the agreement is self-evident. In setting up this example we assumed that a random collection of boxes would show correlation between the three dimensions. A box that is tall is likely to be thick and wide, probably to a greater extent than the correlations which we assumed in the fictitious problem.

It would be possible to set up an example with 20 measurements so that every measurement would involve all three of the box dimensions. A factor analysis of such a system of measurements would not reveal a simple structure, so that no clear bounding planes could be determined. In psychological experiments we must recognize this possibility. Some factor experiments will fail to reveal primary factors. This will necessarily happen if each variable involves as many factors as there are factors in the whole set of variables. There is no guarantee beforehand that this will not happen. On the other hand, it is extremely unlikely that a simple structure will appear with the same identifiable primaries in several independent factor analyses of several test batteries that are given to several populations unless the structure is valid.

The interpretation of the primary factors in Table IV would consist in studying the several test variables that show high saturation on each primary factor. One would try to find what it is that is common to these variables or measures. One might discover that, in general, tallness is a common characteristic of one primary and that the best measurements available for tallness are those which show highest saturation on the factor. We might name it a tallness factor. That would merely imply that the 20 measurements are functions of three linearly independent parameters and that one of them is tentatively named tallness. A box which measures high on that factor would stand higher from the floor than boxes that are low in that factor. Now this identification of the linearly independent parameters in the 20 box measurements would not imply that the tallness factor has some sort of locus in each box, or that the factor is a separate organ or gadget in each box, or that it is some sort of isolated, abstract, statistical box-element that has nothing in common with the rest of the box. It is an independent parameter that is involved in the box measurements. The identification of a primary factor in psychological test experiments is merely a challenge for us to find out what sort of process or attribute it is in terms of known concepts, or else to invent new concepts that fit the experimentally determined primary categories.

Objectivity in Factor Analysis

The attempts to interpret the primary factors have often been criticized because they are not sufficiently objective. Let us be clear about two distinct problems that are involved here. It is one problem to isolate a primary factor and to determine by repeated experiments

that it has some functional uniqueness. That is factor analysis proper. It is another problem to find the psychological or physiological meaning of a functional uniqueness when it is found. That is a matter of interpretation, and consequently it is necessarily subject to debate with conflicting interpretations. We must remember that the interpretation of every scientific experiment is subjective. There is no kind of scientific experiment in which the interpretation rolls out "objectively." If there is such a factor as auditory imagery, for example, just how should the auditory character of the factor produce itself objectively in the statistical analysis and without the subjective interpretation of the investigator? Or just how could tallness in the box population appear objectively in the data? Factor analysis is no exception to all other kinds of scientific experimentation. It is a fortunate circumstance, however, that different interpretations of a primary factor can usually be resolved as questions of fact. New factorial experiments can be made to determine which interpretation is the more plausible.

In naming the primary factors I have taken the explicit policy that it is better to name the factors in terms of well-known concepts such as Number, Space, Verbal, and Memory factors than to name them in some noncommittal way, such as x_1, x_2, and so on. If we name a factor Number, it will provoke experimentation with number tests and with nonnumerical tests, and the experiments are likely to be made in terms of psychological hypotheses that can be sustained or disproved experimentally. In this way we shall advance faster than if the primary factors are left as interesting statistical curiosities. They should be recognized as psychologically challenging experimental effects. I do not expect to be correct in all of the interpretations of the primary factors, and I shall probably have occasion to revise their interpretation repeatedly as more experimental information becomes available. Again this policy is consistent with my desire to make factor analysis a useful psychological tool instead of a mere statistical routine for condensing test scores and correlations.

Factor Analysis and Sampling Theory

In the development of factor analysis it is highly desirable, and eventually it will be imperative, that we find some way of determining the order of magnitude of the sampling errors in the factor loadings, in the test variances on each factor, in the correlations between the primary factors, and in other factorial statistics. In dealing with this problem it has been my belief that it is more

important to find logically correct solutions first and to worry about the sampling errors later, unless the two problems can be solved at the same time. Some statisticians have discarded solutions for which sampling errors could not be determined. An example of the dilemma is that of factoring a large correlation matrix. My first solution to this problem was to determine the principal axes of the configuration of test vectors (**10**). This solution has been stated more elaborately by Hotelling (1) in several papers. But there is practical difficulty in handling a table of correlations for 60 variables. Even with the ingenious iterative procedures of Hotelling the labor for such a large table is prohibitive. The centroid method does enable us to factor a large table of correlations with a reasonable amount of labor, but the nature of the solution is such that sampling theory is not readily applicable to it, if it can be applied at all. The principal axes solution, which has been called the principal component solution by Hotelling, lends itself to treatment in terms of sampling theory. We should like to start with it before rotating the axes to a simple structure if it could be handled without prohibitive cost of labor. This practical difficulty will, no doubt, be resolved. Perhaps machine methods will be found for determining the principal axes solution, and then sampling theory becomes applicable. If such methods can be found, we shall discard the centroid method that we have used so far.

There is another practical solution which is not so elegant, namely, to determine empirically the amount of fluctuation of factorial determinations which are due to sampling. From one fictitious universe of dice throws we might draw several samples and factor each of them independently. The amount of fluctuation which is due to sampling could then be found empirically. The results would have to be interpreted with caution, because the sampling errors would necessarily depend on various characteristics of a particular fictitious universe which would not be duplicated in every practical problem. However, this is one approach that will be useful until factoring methods can be devised that are practical in amount of labor and for which sampling theory is applicable.

In considering the problem of sampling errors we must remember that it is the rotated reference frame of a simple structure that is used for psychological interpretation. Hence, it is the standard errors of factor loadings on the primary factors or primary reference factors that we must have in order to judge the effect of sampling errors on psychological interpretation of a factor experiment. If the standard errors of the projections on the principal axes can be determined,

it should be possible to obtain the standard errors of the projections on the primary axes, for these are linear combinations of the principal axes. But such a solution must take into consideration the fact that the linear combinations are themselves subject to sampling errors. The problem is not simple. In my judgment it is better to proceed with psychological experiments, even though the methods are not yet perfected, and to be ready to reinterpret the results in the light of improvements in analytical methods than to wait for the perfection of method in the abstract. The motivation to improve the analytical methods will derive largely from their need in scientific work that is in progress.

An Experimental Test Battery

In response to a number of requests we have made available for experimental and research purposes a reduced battery of tests which can be given in three sessions of less than two hours each. In preparing a reduced battery for primary mental abilities there are several compromises that must be made in order to make it practical from the standpoint of computational labor for each subject. This work is an attempt to break up a single index of intelligence into component parts which can be psychologically or educationally interpreted. This is a common practice, but the attempt here is to make the component parts of the appraisal fit the primary factors that have been found. In this section we shall consider briefly several of the practical compromises that must be made in such a program.

When a test battery has been given to a group of subjects and the factor analysis completed, it is, in general, possible to appraise an individual as to each of the primary factors, provided that several tests are available with significant saturations on each primary. Each individual primary factor score is determined as a linear combination of all his test scores. If six primary factor scores are to be estimated for each subject, and if each primary score is a function of 16 separate test scores, the computational labor is considerable. Facilitating tables could be prepared for this work, but that has not yet been done. It may be done later, when the tests have become stabilized sufficiently to make it worth while. At present, we are dealing with an experimental edition. As a compromise, we have assembled the tests in groups of two or three for each primary factor. Three tests which have significant saturations in a primary factor can be combined into what we have called a *composite* test score for that factor. In this way we have as many groups of tests as there are primary

factors, and each group gives a composite score. Each individual primary factor score could then be determined as a linear combination of as many composite scores as there are primary factors to be appraised. For the first experimental edition of the tests we have suggested an even simpler arrangement, namely, to treat the composite scores just like ordinary test scores. If a test has a satisfactory validity, the score in the test is used as an estimate of the criterion against which the test has been validated. In the same way we can use each composite score as an estimate of the primary factor which has a significant saturation in each test of the composite. *But these composite tests will be correlated.* Most of the primary factors so far determined have low correlations, but the corresponding composite scores are correlated just like ordinary psychological tests.

There is a misconception about intercorrelations of the tests in a composite. The ideal situation is for the several tests in a composite to have only one primary factor in common, and, hence, the separate tests should, in general, have low correlations. Casual judgment of this problem would perhaps be that the tests within a composite should be highly correlated, but this is not the case. If the several tests within a composite have high intercorrelations, then they may have in common not only the primary factor but also other factors which the composite is not intended to measure. It is better to have the several tests in a composite of low intercorrelation except for the one primary factor which they should have in common. If the saturation of each test is .50 in the common primary, their intercorrelations would be only .25. One should not make snap judgments about the grouping of psychological tests merely by inspecting the raw intercorrelations.

The experimental edition of tests for primary abilities should be used only for research purposes. They are not yet to be regarded as general, service tests for schools and colleges. They are now being given in a number of universities to students majoring in different professions and academic subjects, in order to ascertain whether their factor profiles show differences that can be interpreted.

Individual Laboratory Test Methods

The paper-and-pencil tests which are in general use are subject to considerable chance variation, so that the error variance is quite large. By using individual laboratory tests for each subject the test performances will probably have higher communalities, and the factor methods thereby reveal the primary factors with a smaller element

of uncertainty. A test battery is being assembled with each item on a frame of 35-mm. film. The separate test items are projected on a screen in front of the subject, and his response time is noted with a silent timer. The subject's response for each item is made on one of two keys, one for each hand. The tests have been simplified so that the average response time is usually not more than two or three seconds. Some of the tests will be arranged for voice-key response. The test records obtained in this manner will be factored with the expectation of higher communalities than those in the paper-and-pencil tests. If we should limit ourselves to response time for simple discriminations of various kinds, the time scores would all be in the same units, and the problem of a universal unit of test measurement for factor analysis would be solved.

One of the limitations with practically all psychological tests is that test performance is usually dependent on the subject's motivation to do a particular task and to do it within some restriction of time. The test performance can be no better than the unknown motivation on the occasion of the test. It would be fortunate if the parameters that describe the individual's dynamical system could be determined independently of his motivation to do a particular task as fast as possible on a particular occasion. It may not be impossible to make the appraisal in ways that will avoid this limitation. It would not be sufficient merely to measure the subject's speed factor, if one or more such factors exist, and to discount his other performances in terms of the speed measurements. The motivation would still be unknown. All current tests depend on the implicit assumption that the examiner obtains a uniform degree of motivation. One possibility is to study factorially the individual differences in perceptual Gestalt effects, in the hope that the underlying primary factors may be of some central significance and that their effects are not limited to peripheral or purely sensory detail. Such parameters of individual differences undoubtedly exist, and their determination can be effected without asking the subject to engage in a competitive test performance against time. Whether these parameters will be of general significance for personality traits is an experimental gamble. Other fields should be explored in the hope of eventually obtaining methods of appraising individuals more or less independently of their efforts to perform well on particular testing occasions.

The primary factors that are determined factorially by group test methods should be interpreted psychologically and studied in more rigorous individual experimental form, so that the processes which

characterize the primary factors may be understood. A program of this kind will bring the group testing methods into closer relation with fundamental psychological theory. Testing programs in the schools should then be better founded on psychological theory instead of being merely in the nature of empirical shots at the complex criteria of school performance.

Factor Analysis in Other Sciences

The factor method has been developed for the solution of psychological problems, but it is a general scientific method that can be used in other sciences either in its present form or with adaptations. There is nothing in the logic of factor analysis which limits the method to psychological problems. The old problem of constitutional types is so obviously a factor problem that no student of factor analysis could possibly deal with the problem of body types without seeing it. Customary statistical methods of analyzing such a problem seem futile and fragmentary in comparison with the more powerful factor methods. The method could, no doubt, be applied in psychiatry, where the functional disease entities are not yet determined. Some factor studies have been made by Dr. T. V. Moore in this field, but he has not yet taken advantage of the possibility of a simple structure in his measurements. The disease histories of individuals might reveal factorially how diseases tend to become grouped so that a man who has had a certain disease is more likely to contract certain others and less likely to contract still others. This type of study has been made mainly in connection with body types, but the factor methods have not yet been introduced in that field. One factor study has been proposed to determine whether the scratch tests for allergic patients would reveal primary factors in sensitivity. If such primaries exist, they might be illuminating for theoretical interpretation and also clinically in reducing the number of separate tests which are necessary to cover a wide range of allergic types. Factor methods have been proposed for the comparison of primitive societies in order to determine how many cultural factors are represented in a large number of tribal attributes. A factor study has been made of voting records of a population of districts in order to determine how many primary cultural factors, religion, race, economic status, nationality, and language, are involved in the individual differences among districts on political issues. A factor study is being completed on stock market fluctuations. In the physical sciences it is less likely that the factor methods will find useful application, although opportunities

might be found in some borderline problems. There are similarities between the matrix problems in factor analysis and some matrix problems in modern physics, but the similarities may be only formal.

One important bit of advice may be in order here. The investigator who wants to use factor methods must be competent in the science whose problems he wants to solve. It is not wise for a statistician who knows factor analysis to attempt problems in a science which he has not himself mastered. The most important part of a factorial study is the interpretation of the primary factors that are found and in setting up new factorial experiments to test scientific hypotheses. That is the work of a scientist who first knows the problems of his own science and who uses factorial technique as an aid in the solution of those problems.

My own contributions to factor analysis have been motivated by a desire to solve some fundamental problems in psychology, and consequently I have tried to discourage a tendency to regard the factor method as a self-contained and extraneous statistical technique. Its logic must be dictated by the nature of the scientific problems which it is intended to solve. This paper will have served its main purpose if it helps in some way to bring factor theory away from statistical abstractions and closer to its original purpose in psychological theory.

BIBLIOGRAPHY

1. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. educ. Psychol.*, 1933, **24**, 417–441; 498–520.
2. Kelley, T. L. Essential traits of mental life. Cambridge, Mass.: Harvard Univ. Press, 1935.
3. Spearman, C. Thurstone's work re-worked. *J. educ. Psychol.*, 1939, **30**, 1–16.
4. Thomson, G. H. The factorial analysis of human abilities. *Hum. Factor, Lond.*, 1935, **9**, 180–185.
5. Thomson, G. H. The factorial analysis of human abilities: a rejoinder. *Hum. Factor, Lond.*, 1935, **9**, 358–363.
6. Thomson, G. H. Some points of mathematical technique in the factorial analysis of ability. *J. educ. Psychol.*, 1936, **27**, 37–54.
7. Thomson, G. H. Maximising the specific factors in the analysis of ability. *Brit. J. educ. Psychol.*, 1938, **8**, 255–264.
8. Thomson, G. H. Recent developments of statistical method in psychology—II. *Autumn Occup. Psychol.*, 1938, 1–7.
9. Thomson, G. H. The factorial analysis of human ability. New York: Houghton Mifflin, 1939.
10. Thurstone, L. L. Theory of multiple factors. Ann Arbor: Edwards, 1932.
11. Thurstone, L. L. Unitary abilities. *J. gen. Psychol.*, 1934, **11**, 126–132.

12. Thurstone, L. L. The vectors of mind. Chicago: Univ. Chicago Press, 1935.

13. Thurstone, L. L. Current misuse of the factorial methods. *Psychometrika*, 1937, 2, 73–76.

14. Thurstone, L. L. The perceptual factor. *Psychometrika*, 1938, 3, 1–17.

15. Thurstone, L. L. A new rotational method in factor analysis. *Psychometrika*, 1938, 3, 199–218.

16. Thurstone, L. L. Primary mental abilities. (*Psychometr. Monogr.*, No. 1.) Chicago: Univ. Chicago Press, 1938.

17. Thurstone, L. L. Shifty and mathematical components. *Psychol. Bull.*, 1938, 35, 223–236.

18. Tryon, R. C. Cluster analysis. Ann Arbor: Edwards, 1939.

19. Wilson, E. B., & Worcester, J. The resolution of tests into two general factors. *Proc. nat. Acad. Sci., Wash.*, 1939, **25**, 20–25.

20. Woodrow, H. Factors in improvement with practice. *J. Psychol.*, 1939, 7, 55–70.

21. Woodrow, H., & Wilson, L. A. A simple procedure for approximate factor analysis. *Psychometrika*, 1936, 1, 245–258.

22. Wright, R. A factor analysis of the original Stanford-Binet scale. *Psychometrika*, 1939, 4, 209–220.

23. Zubin, J. Review of *Primary mental abilities*. *Psychol. Bull.*, 1939, 36, 204–208.

STUDIES OF SIMIAN INTELLIGENCE FROM THE UNIVERSITY OF LIÉGE

BY K. S. LASHLEY
Harvard University

In a series of papers published during the past seven years, Professor L. Verlaine and his collaborators at the University of Liége have reported a large number of experiments dealing with the intelligence of macaques. The experiments have been cleverly devised to test the capacity of the animals to generalize from specific visual experiences and have led to important conclusions concerning the primary role of generalization in perception. The results of the later studies indicate for the macaque a level of capacity above that currently ascribed either to monkeys or anthropoid apes. If valid, the studies are among the most important recent contributions to comparative psychology.

The journals in which the work has been reported are rather inaccessible in America, and the published abstracts have failed to indicate the character of the data. Since much similar work is being done in this country with results often contradictory to those of the group at Liége, a critical review of the studies carried out there seems timely. They have dealt chiefly with the capacity of the monkey to abstract general properties of the stimulus from a number of specific experiences, and have ranged from the familiar tests for transposition in a series varying in a single dimension to the requirement of discrimination based upon such difficult conceptual abstractions as that of a "mammal." In describing the work I shall not be concerned with the theoretical interpretations which the authors have formulated, but only with the nature and validity of the experimental results.

Methods

In general, the procedures were those of the familiar discrimination experiment in which the animal chooses between two or more objects and is rewarded for taking those which the experimenter wishes him to select. The methods are not reported in detail, so that it is not possible to discover just how the stimuli were presented in the various experiments. In the earlier work, figures cut from paper were used as stimuli and were apparently placed on the table before the animal, with or without a reward (meal worm or bit of banana) under them. In later experiments a small metal box lid was placed

beside each stimulus object, with the reward under that accompanying the positive stimulus. The animal was chained in a shallow pen or box and the stimuli arranged on a low table attached to one side of the pen (figured in **17**). More frequently, apparently, the stimuli were arranged on a cardboard tray, which was held in the hand of the experimenter while the monkey made his choice of the objects presented (figured in **18**).

Three methods of choice were used. (1) *Simultaneous presentation* of stimuli: Two to six objects are arranged on the tray before the animal, half the objects, as three different triangles, concealing food and half, as three nontriangular figures, without food. (2) *Successive presentation:* A single object on the tray is held out to the animal. He must take it or refuse it, according to its visual properties. (3) *Choice from model:* The experimenter arranges two objects on the tray with food under one of them. He places the tray before the monkey with his left hand, at the same time presenting with his right hand, "above and between" the two objects on the tray, a third, similar to that concealing the reward. The animal must choose the object resembling this 'model.'

It should be noted that only the first of these methods permits of any adequate control of extraneous cues from the experimenter.

The training differed from that employed by other investigators in that identical stimuli were not used in successive trials, but only stimuli having some common property. Thus, in one experiment (**14**), 54 triangles differing in size, shape, or brightness and 81 nontriangular figures differing similarly were prepared. These were presented in any of four orientations, giving 538 different figures. In training, a single figure was offered at each trial, all triangles with reward, others without, and the same figure was not used twice during the experiment.

This method may have a very decided advantage over training with constant stimuli, since it may induce the animal to abstract the property common to the varying stimuli and so avoid a more rigid association with a specific object. In my experience, however, a continuously changing problem is likely to cause a complete disregard of the experimental stimuli in favor of extraneous cues.

Animal Subjects

Most of the experiments were made with a male Bonnet monkey *(Macacus sinicus)*, named Coco, about two and one-half years old at the beginning of the work. He had been kept as a pet for a year,

and his previous history was unknown. In a few experiments a second monkey of the same species and a female rhesus, Venus, born in captivity, were also used.

The first experiments with Coco were performed by Verlaine and Gallis (**20**). The monkey learned, after somewhat prolonged training, to open a box by inserting a nail in a small hole. He then readily used other small objects as substitutes for the nail, even breaking a branch from a plant and a straw from a broom for this purpose. He transferred the use of a stick from opening the box to pushing food from a tube, after some random activity. Coco also succeeded in tests requiring the use of a box as a ladder and in other motor performances which, though unusual for a macaque, are well within the capacities of the Cebus monkeys studied by other investigators.

The investigations were next directed to an analysis of the capacity of the monkey to distinguish and generalize the properties of geometrical and other figures presented visually or tactually. The earlier studies of this character were carried out by Miss Tellier. More complicated problems were later devised and used by Verlaine.

Reactions to Unidimensional Series

A long list of experiments on the basis of reaction to quantitatively variable stimuli is presented (**3, 6, 8**). The variables studied were size, height of blocks, dimensions of angles, spacing of blocks, luminous intensities, color in a series of mixtures ranging from red through purples to blue, visual distance, speed of movement, weight (**6**) and roughness of sandpapers (**3**). The method was to train the monkey to choose the shorter of two rectangles or the more distant of two blocks, for example; next, to test for transfer of reactions to other magnitudes in the scale; then, to train for the selection of a specific member of the series, opposed to all others.

With color, speed of rotation of a vane, and spaces between blocks the results were rather inconclusive. With all the remaining series the animals tended to choose on a relational basis, and reaction on an absolute basis could be established only with difficulty, after considerable training. The results are in harmony with studies of transposition made by other investigators and include a wider range of variables.

A number of interesting incidental observations are included, of which I have space to mention only one. Venus, previously trained to choose the shorter of two rectangles, was taught to choose one 12 cm. in length when opposed to any other from 6 to 20 cm.

She was then tested with pairs not including the 12-cm. size. When 10- or 14-cm. lengths were opposed to others not very different from them, she chose the one nearest 12 cm. in length, but, with rather great differences between the members of the pair, she regularly chose the shorter. Both relational and absolute properties were thus effective in the same series of tests (**8**).

There was a little evidence (**5**) that, in learning a new series, the animal tends to choose on a relative or absolute basis, in accord with his training on other series. This evidence is discussed by Verlaine (**14**, pp. 32 ff.), who concludes that the animal generalizes the nature of his attack of the problem, irrespective of the material. This conclusion is based largely upon two critical trials with Venus and ignores the fact that immediately following trials gave opposite results (**8**, p. 48). The conclusion is scarcely established by the data.

Reactions to Number

The experiments of Gallis (**1**) dealing with counting by monkeys are remarkable only for the neglect of every experimental control. Tellier (**8**) prepared cards with different numbers of dots, from 2 to 11. Five cards, each with a different arrangement of dots, were used for each number. Coco was trained to choose three from 2 and 4, five from 4 and 6 dots, etc., up to ten from 11. He learned all of these problems. His errors were said to resemble those made by a person judging the relative concentration of the dots without counting them. No tests were run to determine whether Coco finally learned by actual counting or only by recognizing the specific patterns of each set of five cards, but in later discussions it is taken for granted that Coco could count.

The Generalization of Form

Tellier (**7**) reports the earliest experiments dealing with the abstraction of the form of visual figures. Coco and Venus were trained with nine triangles, differing in shape, size, or brightness, as positive stimuli and nine different squares and rectangles as negative, presented in pairs in chance combination. The monkeys learned to choose all of the triangles after 139 and 171 errors, respectively. The method was then changed to the simultaneous presentation of three positive and three negative stimuli. When accurate discrimination was established under this condition, various concurrent habits were established: *e.g.* to choose any triangle and also a particular parallelogram from among a variety of other figures; to choose a white circle, a black square, and a gray trapezoid when presented together

with three other figures of whatever sort. These experiments illustrate the monkey's ability to recognize similarities of form and to acquire several habits simultaneously. They do not go beyond the results of other investigators.

At this point in the experiments the method of choice from a model was developed, and most of the later work has been done with this method or with the successive presentation of single figures. Verlaine (**14, 19**) later took up the problem of generalization of geometrical forms, using first the method of successive presentation. Coco was trained with figures, selected from 216 different triangles and 324 nontriangles cut out of paper, to accept any triangle and reject any other figure when offered alone. No figure was presented twice during the training. He learned the problem in about 200 trials. He was then given the following succession of tests:

(1) To recognize the figures when all but a narrow margin, one side with the two including angles, was hidden by a card. This amounted to recognizing whether or not the two edges interrupted by the covering card would intersect to form a triangle, if prolonged. Coco grasped this problem at once, making only 10 mistakes in 200 trials.

(2) The paper figures were cut into two pieces, and the two pieces, separated by an interval of 15 mm., presented. Coco must react positively if the two pieces united would form a triangle, negatively if they formed any other figure. This problem was solved without error.

(3) The parts of the figures were arranged on the tray with their normal relations reversed or with one part inverted. Coco must reassemble them 'mentally' and accept only resultant triangles. With this problem he made the first error on the fifteenth trial and only three errors in 30 trials.

(4) Two fragments of triangles were placed in random position on the tray. They were to be selected if they would make a perfect triangle when combined; rejected if they could not be fitted together to form a single triangle. Coco had to determine this by inspection, without touching the fragments. After five errors in the first 10 trials he made above 80% correct.

(5) Variously shaped pieces were cut from the middle of triangular papers. At each trial one of these triangles was presented with a small piece of paper beside it. If the small piece would fit the hole, the stimulus was positive; otherwise it was negative. Coco's score was 5–7–7–8–9–10 correct in successive 10's of trials.

The method was then changed to choice from a model and line drawings, black on white paper, substituted for the forms cut out of paper.

(6) Two incomplete figures, *e.g.* a triangle with two corners omitted and a circle lacking a quarter segment, were presented on the tray. At

the same time a third figure, *e.g.* a square, was shown above them. Coco must identify the incomplete figures and choose a triangle or nontriangle according to the model; the part circle when the square was shown, the incomplete triangle when a triangle was shown, etc. He grasped this problem at once and made a perfect score with a variety of figures.

After various other experiments of this type, giving similar results, the following problems were set.

(7) Choice from model: Two incomplete triangles were presented on the tray. A card on which was drawn the missing part of one of the two was shown as a model. Coco must choose that one of the two figures which could be combined with the model to form a perfect triangle. He solved this problem without an error. He solved it, also, almost immediately when it was presented, with the necessary modifications, by the methods of simultaneous and successive presentation.

(8) Cards on which were drawn three or four lines and two, three, or four short-sided angles, in random arrangement, were presented in pairs. Coco must choose the cards bearing three lines and three angles and reject any other combination which would not form a perfect triangle, with nothing left over. Coco comprehended the problem at once, making his first error on the nineteenth trial and 19 out of 20 correct.

The most remarkable thing about these and subsequent experiments is not that the animal could recognize the properties and relations of the triangles, but that he should discover the purpose of each experiment without preliminary training, shifting immediately from the choice of the more complete of two imperfect triangles to the choice of the one which would complete the model, etc. However, this is among the least of his achievements.

Other General Conceptions

Tellier (7) trained Coco to choose pictures of animals and reject pictures of plants. She used colored pictures of 26 animals, ranging from butterfly to lion, and 18 pictures of leaves and flowers. At each trial a different pair, animal and plant, was presented. Coco learned the series to 10 errorless trials in 70 trials, with 25 errors. (This means that each picture was not seen more than three times.) He was next trained to choose quadrupeds and reject birds. For this he required 446 trials with 202 errors. Tested immediately with tracings in outline of the first series of colored pictures, he made only two errors, rejecting plants and selecting both quadrupeds and birds, in spite of the immediately preceding training to reject the birds.

Verlaine (**14**, **19**) later continued experiments with this material, using the method of choice from model. The order of succession of

experiments is not clear, but, apparently, all of the experiments dealing with generalization of triangles, reported above, had intervened since Tellier's use of the pictures.

(9) Coco was required to choose the picture of an animal or of a plant according to the type of the model. He remembered perfectly his earlier training, grasped the problem at once, scoring 8–9–8–10 correct in successive 10's of trials. He made equally good scores by the methods of simultaneous and successive presentation.

(10) The pictures of animals and plants were cut in two. Coco must choose part of an animal if an animal were shown as a model; part of a plant, if a plant were shown. Thus an elephant was the indicator for choice of the tail of a cock rather than leaves and flowers of a pansy; a leaf of marigold called for choice of a red tulip and rejection of a headless cheetah. Coco grasped the problem at once and made no error with 11 sets of figures.

(11) Actual animals and plants were substituted for the pictures, and Coco was required to distinguish between parts of real animals and plants. Ten sets of objects were used, such as the following:

Positive	*Model*	*Negative*
picture of a dog	hornet (dead?)	two dried raisins
bee on a pin	wasp comb with pupae	piece of potato
pictured head of a chicken	live fish in bowl	picture of tulip
chicken feather	picture of a duck	potato peeling
lock of brown hair	part of a white feather	pinch of tea

Coco made five errors in the first 10 trials, three in the first repetition, then an errorless score. This score should be compared with the 446 trials and 202 errors required to learn quadrupeds and birds in Tellier's experiment. It should be noted, also, that immediate success in this problem requires that Coco should have learned from the pictures of leaves and flowers not only the general properties of the pictures, but the substantial structure of plants, so that he could immediately identify a potato peeling as having properties like a pictured leaf and flower, or recognize a wasp's nest as an animal product.

A number of additional experiments of this type are reported which it seems useless to review in detail. Whatever the problem set, Coco grasped it at once and chose correctly. One more example will suffice:

(12) Two pairs of fragments of animals were presented on the tray. One pair was from the same animal, the other pair from different animals. Coco must choose the perfect or imperfect animal according to the character of the model. Thus a fish with the tail of a bird, presented as the model, calls for selection of a wasp with the head of a hornet and rejection of the head and body of a hornet. Coco grasped this problem at the first presentation and made no error with 10 sets of

objects. In other tests he correctly distinguished the combined parts of birds of one species from combinations of parts from different species.

Other publications deal with the conception of substance, of the living and nonliving, and with the utilization of former associations by contiguity.

(13) To test his conception of substance or material Coco was confronted with 13 pairs of objects of unlike material, such as metal, cloth, pottery, etc. (12). Of these he must choose according to the material of the model, thus:

Positive	*Model*	*Negative*
forceps	nail	pencil
white cloth	hat	plate of zinc
glass tube	fragment of china	piece of sugar

With no previous specific training, Coco understood this problem immediately and made errors only on the second and twelfth trials. Generalization of the animate and inanimate was equally prompt and certain (11).

Finally, an experiment was devised to determine whether Coco could utilize a number of incidental associations formed during his life in the laboratory (16). Had he noticed that the students put sugar in their tea cups but not in the saucers; that meal worms were fed on bran and not on tea leaves; that photographic films were hung to dry with clips and not with screws; that the roots, but not the stems, of plants are imbedded in the soil? He was offered pairs of objects and shown a model which he should have associated at some time with one only of the objects, thus:

Positive	*Model*	*Negative*
cup	sugar	saucer
clip	roll of film	screw
cage	bird	box
pinch of soil	earthworm	bran

In 50 different tests of this sort Coco made only five mistakes. Verlaine analyzes the mistakes and decides that most of them are probably not mistakes. Thus, confronted with paper and a sheet of metal, with scissors as model, Coco chose the sheet of metal. This error is excused by the statement that he had doubtless seen sheet metal as well as paper cut with scissors.

How are we to evaluate these reports? Coco, at the age of four, succeeds in picture completion tests at the five- to seven-year human level; his classification of animals and plants is not much inferior to that of the early naturalists; his knowledge of ornithology exceeds that of the reviewer; he grasps the purpose of every experiment most often without a single error. If the experiments are valid, they

justify Verlaine's conclusion that the monkey thinks just as man does, though not so much (20).

A number of other investigators have trained individual monkeys for periods as long as that covered by the experiments with Coco. They have not tried tests like those of Verlaine for the simple reason that their animals have failed in problems so much more simple that it seemed hopeless to complicate them further. The few experiments reported in the literature which have sought to go beyond generalization at the direct perceptual level have given negative results.

If Verlaine's experiments are valid, he has, then, either discovered a genius among monkeys or developed unusual capacities through his unique methods of training. The early records of Coco, in the hands of Tellier, are not markedly superior to those of the other Bonnet or Bengal monkeys which she trained nor are his achievements, prior to the training in choice from model, unusual for a macaque. There is little reason to ascribe the later results to the fortunate acquisition of a very superior animal.

Verlaine's experiments are a masterpiece of planning. In successive tests he leads from simple to more and more difficult concepts, so that success in each experiment generally might give a clue to the solution of the next problem. If the monkey really is capable of acquiring such abstract concepts, then no better pedagogical system could have been devised than that which Verlaine has developed. Nevertheless, the systematic arrangement of the experiments will not account for all of the cases in which Coco instantly grasped a difficult concept.

In Tellier's earlier experiments, for example, Coco was trained to identify wooden blocks of various shapes: cone, cube, etc. Tested with line drawings of these objects, he failed to transfer and required 104 trials with 43 errors before he learned to distinguish the line drawings. This is as bad as his learning scores with totally new figures and suggests that the line drawings were in no way associated with the solid objects, nor was there any further training to establish such an association. Nevertheless, a little later he immediately identified the tracings made from colored pictures, and in Verlaine's experiments, with no further practice in recognition of similarities of pictures and objects, he not only identified objects from their pictures, but transferred the general concepts of animal and plant, derived from pictures of mammals, birds, leaves, and flowers, to such objects as the wing of a wasp and a bit of skin of an onion. Coco's rapid acquisition

of difficult concepts under such circumstances cannot, then, be ascribed to uniquely efficient methods of training.

There remains for consideration the possibility that none of Coco's more remarkable achievements are genuine. By contrast with the ingenuity and insight shown by the experimenters in planning the tests, the actual technique of experimentation seems incredibly naive. As controls of extraneous cues Tellier mentions only placing meal worms under the negative as well as positive stimuli, and in one case, where the problem had been difficult (**8**), Coco failed to discriminate under this condition. Verlaine gives only a brief and unsatisfactory account of controls (**14**, **19**). Odor, perhaps with monkeys the least probable of extraneous cues, was regularly controlled. Precautions against cues from the experimenter are described as follows: "I have not spoken to him (Coco) or made any sort of sound during any of the tests to which he has been submitted." "I have always taken care to adopt an absolutely neutral attitude and to place the tray before him in the same way, so that he should not gain from any of my gestures or attitudes . . ." a cue to the position of the positive stimulus. "I have even presented the tray with my back turned to him and have intentionally placed the negative stimulus in front of him."

For the method of successive presentation the following controls are reported. The tray was offered more slowly with the positive, more quickly with the negative stimulus. Coco's reactions were not disturbed. "This did not give the impression that the animal was easy to influence." A single test with a more mechanical control is described. The positive or negative object was placed alone on a board from which a string led within Coco's reach. Coco pulled in the board when it bore the positive stimulus and failed to do so with the negative in 41 of 50 trials. In these tests the author prepared the stimuli, "as usual" with his back turned to the monkey, turned and placed the tray on the board, "sans regarder moi-même ce que ferait Coco."

To test "telepathy," strangers without knowledge of the purpose of the experiment tested Coco. The nature of the tasks is not stated, except that they were among the more difficult, nor is it stated whether or not Verlaine and other familiar experimenters were present during the tests.

Coco's more remarkable successes were attained under the conditions of selection from model or successive presentation. These are the situations in which the animal may most easily discover significant

cues in the behavior of the experimenter, since in the first case the model is held in the experimenter's hand and in the second the animal need only get a cue to take or refrain from taking. The most elementary precautions of screening the experimenter from the animal during the preparation and presentation of the stimuli were completely neglected. Tracing the history of the experiments, the most probable inference is that Coco first began to detect extraneous cues during the later experiments of Tellier, at about the time when he was trained in choice from model. Up to this time most of the tests had involved simultaneous presentation of several negative and positive stimuli, a condition under which directional cues are difficult to convey. Thereafter the methods were limited to those under which Coco was most successful, and it is scarcely a coincidence that these were also the methods which give the best opportunity for effective extraneous cues.

The studies have required an enormous amount of work. They show an extraordinary ingenuity in devising successions of experiments leading to more and more general relations between the objects to be discriminated, successions which might well lead an intelligent subject to discover the general attributes of the stimuli. Nevertheless, in view of the lack of any adequate control of extraneous cues and the failure of other investigators to obtain remotely similar results, the achievements of the monkey of Liége must be classed with those of the horses of Elberfeld and the dog of Mannheim.

BIBLIOGRAPHY

1. Gallis, P. Les animaux savent-ils compter? *Bull. Soc. Sci. Liége,* 1932, 1re année, 82–84.
2. Tellier, M. Reconnaissance par le toucher d'objets connus par la vue, chez le macaque. *Bull. Soc. Sci. Liége,* 1932, 1re année, 114–117.
3. Tellier, M. Le sens du toucher et la généralisation, le relatif et l'absolu, chez le macaque. *Bull. Soc. Sci. Liége,* 1932, 1re année, 192–196.
4. Tellier, M. Le macaque saisit-il le rapport logique? *Bull. Soc. Sci. Liége,* 1932, 1re année, 227–231.
5. Tellier, M. Le choix d'après le modèle, chez le macaque. *Bull. Soc. Sci. Liége,* 1933, 2me année, 41–45.
6. Tellier, M. Le sens du toucher et la généralisation. La discrimination du poids des objets chez le macaque. *Bull. Soc. Sci. Liége,* 1933, 2me année, 138–142.
7. Tellier, M. L'intelligence des singes inférieurs. La vision des formes et la généralisation. *Mém. Soc. Sci. Liége,* 1934, **19**, No. 3, 1–76.
8. Tellier, M. L'intelligence des singes inférieurs. II. Le relatif et l'absolu. *Mém. Soc. Sci. Liége,* 1935, **20**, No. 1, 1–64.

9. Verlaine, L. Le relatif et l'absolu chez le macaque. *Bull. Soc. Sci. Liége*, 1935, 4me année, 90–96.

10. Verlaine, L. À propos du syncrétisme. La perception d'un objet chez le macaque. *Bull. Soc. Sci. Liége*, 1935, 4me année, 132–137.

11. Verlaine, L. La perception chez le macaque. Le vivant et le non vivant. *Bull. Soc. Sci. Liége*, 1935, 4me année, 137–140.

12. Verlaine, L. De la connaissance chez le macaque. La substance. *Bull. Soc. Sci. Liége*, 1935, 4me année, *239–242*.

13. Verlaine, L. Le charactère analytique de la perception chez le macaque. *Bull. Acad. Belg. Cl. Sci.*, 1935, **21**, 744–758; 800–816.

14. Verlaine, L. La vision des formes chez le macaque. Syncrétism, analyse et synthèse. *Mém. Acad. Sci. Belg.*, 1935, **14**, Ser. 2, No. 5, 1–85.

15. Verlaine, L. L'analyse et la synthèse dans la perception des formes chez le macaque. *Ann. Soc. Zool. Belg.*, 1935, **66**, 57–66.

16. Verlaine, L. Les associations par contiguité chez le macaque. *J. Psychol. norm. path.*, 1935, **32**, 719–730.

17. Verlaine, L. Histoire naturelle de la connaissance chez le singe inférieur: I. Le concret. *Actualités sci. industr.*, 1935, No. 215, 1–55.

18. Verlaine, L. Histoire naturelle de la connaissance chez le singe inférieur: II. Le syncrétique. *Actualités sci. industr.*, 1936, No. 320, 1–63.

19. Verlaine, L. Histoire naturelle de la connaissance chez le singe inférieur: III. L'abstrait. *Actualités sci. industr.*, 1936, No. 360, 1–58.

20. Verlaine, L., & Gallis, P. L'intelligence des singes inférieurs. *Mém. Soc. Sci. Liége*, 1932, **17**, No. 12, 1–48.

BOOK REVIEWS

MOORE, T. V. Cognitive psychology. New York: Lippincott, 1939. Pp. viii+636.

Those who know that Professor Moore is a monk of the order of St. Benedict, that he is a psychologist at the Catholic University of America, that his book displays prominently the ***nihil obstats*** of the church authorities, need not be told that this volume is an orthodox scholastic production. The doctrine of the book is revealed in the following quotations:

"One cannot account for the facts of consciousness without a psyche, that is to say, a unit being that is conscious. . . But if there is experience, there must also be someone who experiences; if there is adjustment, there must be someone who adjusts. That which experiences and adjusts is the psyche" (pp. 3, 4).

"The intellect is a power of the soul, a truly existing substantial being. The soul has powers immersed in the organs of sense; but the intellect, which is also a power of the soul, involves the activity of no bodily organ and hence cannot be destroyed with the destruction of the body, but is by necessity immortal" (pp. 119, 121).

It remains for the reviewer to point out that this is a very competent textbook, excellently and sometimes poetically written. In referring to Gamper's *Mittelhirnwesen,* Professor Moore writes that this little creature "may possibly have smiled because it was delighted with the ripples of consciousness that passed over the shallow waters of its mind" (p. 70). Another creditable feature of the book is that the discussion of cognitive phenomena is well oriented with historical material. Here an interesting item is the criticism of Hume, who "in the seclusion of a retired spot, without library, laboratory, or personal contacts, worked out his theory of the human mind." Professor Moore comments: "From such attempts it was not likely that anything of permanent value could be developed" (p. 177). Helmholtz, on the other hand, is lauded for standing at strategic points on scholastic ground. "He (Helmholtz) expresses the Thomistic concept that certain ideas can be termed innate only in the sense that the capacity to acquire them is innate" (p. 190). Another interesting item is that "the configurational psychology of today is essentially a stepping stone from sensationalism to scholasticism" (p. 204).

Although it must be a foregone conclusion that all factual evidences coördinate with the original scholastic premises, Professor Moore profusely documents his book with experimental and clinical findings. For example, in the interest of a psyche independent of the nervous system he suggestively covers the findings of Dandy, Gardner, Howe, Alford, Penfield and Evans, etc.

The reviewer strongly recommends this book for every student's reading list. In the first place, whatever may be the influencing conditions, scholastic ideas are increasingly becoming items in our intellectual culture. A recent book advertisement not ineptly states:

"Much of University of Chicago President Robert M. Hutchins' theory of education (now being put into practice at St. John's College, Annapolis, Md., and watched intently by educators all over the country) is based on the system of philosophy evolved by St. Thomas. Elsewhere too there is a revival of interest in St. Thomas' writings which to many seem to hold a special significance for our time."

It is well, then, that the student know what these ideas are.

In the second place, in the interest of a naturalistic and objective development of psychology—which is likewise a prominent feature of our culture—it is important for students to assess psychological doctrines and points of view which are ostensibly derived from observation, but which have actually originated in historical traditions. Students of psychology, perhaps more than others, should know the origin and nature of doctrines in order to evaluate them. A reading of Professor Moore's book should help to stimulate an effective critical attitude, since it frankly avows psychic doctrines which the nonscholastic treatises imply but effectively conceal. Among the ideas the student will glean, the following may be mentioned: Experimental and clinical means do not justify every postulational end; experimental evidence may be offered as support for what one believes on faith; unless one powerfully cleaves to natural events, one is likely to interpret and use them as something entirely different.

J. R. Kantor.

Indiana University.

Eddington, Sir A. The philosophy of physical science. New York: Macmillan; Cambridge, England: University Press, 1939. Pp. ix+230.

It is the fate of the mystic that we know even before he starts on his journey what goal he will reach. Whether he wings his unerring way through an unresisting ether or plods with ponderous step across a stony terrain, he inevitably reaches the cosmos of his consciousness—a world without center or circumference. Yet it is the fortune of the mystic—if he be as clever as Eddington—that his peregrinations are not uninteresting. How does he refurbish the old terminological and rhetorical equipment? What modern scientific findings does he employ to bring about an anciently concocted dénoûment?

In this small volume, constituting the substance of the Trinity College, Cambridge, Tarner lectures for 1938, Eddington chooses a different course from that in *The nature of the physical world.* Here he is not interested in the nature of the world of physics, but rather in establishing a mystic philosophy on the basis of an epistemological argument, an argument having two phases.

As a student of physical science he begins with the results of selectivity in the theories of quanta and relativity. The obvious fact that science is selective and constructional he pushes to the point of subjectivity and therefore makes physical propositions, if not physical phenomena, into exclusive creations of the scientist's mentality. The argument here is that the physical universe (that is, the universe of physics) is a mathematical structure. The subjectivistic argument is

supported by the capitalization of ignorance. We ought to have or want to have absolute knowledge, but we don't have it—therefore subjectivity. Since we cannot know at once the precise position and velocity of an electron, then all is indeterminate. Selection is transformed into personal creation, and "all the fundamental laws and constants of physics can be deduced unambiguously from *a priori* considerations" (p. 62). In this phase of Eddington's epistemology the scientist is pictured as operating in a vacuum. He allows for no gradual and cumulative development of scientific results (data and ideas) from contacts with things. Eddington is still troubled by the question whether an external world exists, though his problems really concern the mass, magnitude, luminescence, density, and interrelations of astronomical bodies. Because physical and astronomical phenomena allow for tremendous extrapolatory abstraction, need we be blinded to the palpable fact that science and epistemology cannot lose their character of human enterprise designed to attain approximate statements concerning particular problems?

Next, Eddington plunges into the morass of mind, asserting that "the starting point of physical science is knowledge of the group-structure of a set of sensations in consciousness" (p. 148). Thus he attains to the mysticism which for many centuries has been the faith of believers, though they have used other materials and other proofs to implement the common faith. Here our author luxuriates in his ignorance of those scientific developments which make unnecessary any commerce with such vestiges of medieval thought as sensations, consciousness, and mind as knowing spiritual elements. Though we cannot blame a mathematical physicist for lacking acquaintance with modern objective psychology, neither are we obliged to accept his construction of the philosophy of physics. In view of the fact that a philosopher, as against a scientist, of physics must be concerned with the interrelations of the sciences, can we escape the conclusion that an acquaintance with objective psychology is necessary for the physical philosopher? Not the facts observed, not the methods of observation make the philosophy of physics mystical, but rather "a mystic source welling up in our nature" (p. 222).

Though it was easily predicted that in this volume, as in his others, Eddington would turn out to be the same mystic interpreter of the universe, the reader profits from examining it. He learns the dark and devious procedure whereby a scientist of attainment is attracted and baffled by absolutes of various sorts. He may learn, too, that the penalty for deprecating physical science because it can only (*sic*) yield relations and probabilities instead of the absolute means to know the universe is the conclusion that "in the age of reason, faith yet remains supreme; for reason is one of the articles of faith" (p. 222).

J. R. Kantor.

Indiana University.

Wechsler, D. The measurement of adult intelligence. Baltimore: Williams & Wilkins, 1939. Pp. ix+229.

The system of tests described in this work is called the *Bellevue Intelligence Tests,* and the materials required for its use are handled by the Psychological Corporation.

The book opens with a ten-page chapter on the nature of intelligence, which is defined as "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment." No question is raised by the author concerning the measurability of intelligence as thus defined.

The need for a test system intended specifically for study of adult intelligence as opposed to juvenile intelligence is stated in a way which must strike a responsive chord in every clinical examiner who has had occasion to present the Binet scale to a mature subject of average or superior achievement. It is pointed out, also, that the inconstancy of the IQ (as determined by the currently accepted method) is an inevitable consequence of "the assumption of the linear relationship between M.A. and C.A." The rapidity of mental development in younger children is shown very neatly in a mental growth curve based on Bellevue test results obtained from children of the ages seven to seventeen.

Two types of mental deterioration are recognized. In addition to the pathological impairment following brain lesion or prolonged mental illness, there is also a "normal mental deterioration" which occurs after maturity with the natural increase of age. Every human capacity, after attaining a maximum, begins an immediate decline which is at first very slow, but which increases appreciably after a while. This is shown graphically, the results of the Bellevue tests indicating that these test-scores decline a little more rapidly than does vital capacity as measured by lung capacity. Some abilities decline more rapidly than others, and specifically the ability used in the General Information test holds up much better than the ability called for by Memory Span for Digits; but no ability remains constant, and the effect of using a single denominator for calculating the IQ of adults is to "destroy the very basic concept of the IQ." For a valid evaluation of an individual's brightness, one must compare his mental ability with that of the average individual of his own age.

In spite of his opposition to the IQ as we know it, the author considers "Intelligence Quotient" too happy a term to be discarded. In his standardization of the Bellevue tests he presents the norms, as determined by standard deviation procedure, in tables by means of which the weighted scores can be converted directly into what he calls the "IQ." This use of the term, as applied to something entirely different from what we understand by it, is likely to be misleading.

The strongest feature of the system is the selection of cases for standardization. The norms are based upon scores obtained from 1750 subjects ranging in age from seven to seventy, selected out of 3500 subjects to whom the tests had been presented, the selection being a sampling based upon the occupational distribution of the country's adult white population, as indicated by the 1930 census. The adult subjects were divided into age groups by five-year intervals, the number of cases in a group ranging from 50 subjects in the later fifties to 195 subjects in the later twenties. The collection of such a mass of data is a marvelous achievement.

The scale consists of ten test units (plus one alternate), each unit hav-

ing a wide range of discriminative capacity. Correlations and intercorrelations have been calculated for all possible combinations, and the units are ranked in value according to their correlation with the scale as a whole. The scores have been so weighted as to make the numerical values approximately uniform for the eleven units. Three tables of norms are based upon these weighted scores: one for ten tests, one for five language tests, and one for five performance tests. The student who is not too scrupulous about manipulation of results will be able to obtain approximate ratings from other combinations.

Most of the tests are modified forms of tests already in use. The strongest unit of the scale is "Similarities," a series of 12 pairs beginning with *orange—banana* and ending with *fly—tree.*

Following Binet, Dr. Wechsler has apparently given more careful attention to criteria for standardization than to criteria for selection of test items. He includes Memory Span for Digits, although acknowledging it to be a weak test as a measure of intelligence. Furthermore, he combines digits forwards and digits backwards under one score, allowing the same credit for five digits forwards as for five digits backwards. He uses also the Symbol-Digit test, observing that the scores show a marked decline with advancing age, but making no mention of visual strain as a possible explanation of this decline. In the tests called "Picture Arrangement" and "Picture Completion," he has used pictures which are too small and too sketchy to be easily understood. It appears that the value of a test, for him, is statistical rather than psychological.

The text is carelessly written. "Like" is used as a conjunction, and "where" as a relative pronoun. Throughout the book the author refers to himself as "we," as if the work were a joint product. His irregular use of the "editorial we" is at times misleading, especially when he uses "we" as including (apparently) the small group of students who assisted in making the examinations or the larger group of psychologists who may be expected to use the test in the future.

Some of the test materials are contained in six pasteboard boxes, each of which has the name *Bellevue Intelligence Test* blazoned on the cover. Before placing these boxes on the table in full view of the subject, the examiner would do well to cover the offensive labels with something in code.

Clinical examiners who are relatively satisfied with some form of the Binet-Simon scale for examination of children will welcome a test which is intrinsically better adapted to adult interests and which is adequately standardized for adults. To this extent, the Bellevue scale will meet a need that has been keenly felt for more than twenty years. It does not, however, show any such advance over the Binet scale as might be expected as a result of our collective experience in the use of tests for a full generation.

Students engaged in the development of tests for individual study can learn much from Dr. Wechsler's plan of standardization. The table by means of which a score can be converted directly into the IQ (or its equivalent, preferably under some other name) marks a very important advance over the current method of deriving the IQ. If all tests were

thus standardized, the public demand for a one-figure rating could be satisfied honestly and without distortion of results.

But the "mental age," although it lends itself to gross abuse, need not be wholly discarded. Many students, with full appreciation of Dr. Wechsler's contribution to higher standards of accuracy in evaluation of test results, will be disappointed because he has made it difficult to derive a satisfactory "mental age" from the scores of his tests. So long as this way of expressing results is so nearly universal, the "mental age" is indispensable as a common denominator. It is a mistake for any author of tests to assume that others will wish to adopt his system as a whole and use it exactly as he himself uses it. Some of us prefer a flexible system to any composite scale. If Dr. Wechsler had included a table of true "mental age" norms for the ages seven to seventeen, not for his three scales, but for each individual unit, the Bellevue tests would be incomparably more useful and would presumably see much wider service.

A full-size textbook is too large to be used conveniently as a manual of instructions for test presentation. The third part of this book, containing the instructions and norms, might better have been separated from the reference material in the first two parts. For the convenience of the examiner it is to be hoped that the author will publish additionally a compact booklet or pamphlet containing only what is needed for the daily use of the tests. It is strongly urged that he offer norms (for each unit individually) in the form of simple decile scores for each age group, unweighted and wholly free from mathematical manipulation. He has in his possession much material which would be invaluable to clinical examiners.

GRACE H. KENT.

Danvers State Hospital.

MARGADANT, S. V. Eine tiefenpsychologische Grundlage zur Klages'schen Graphologie. Amsterdam: N. V. Noord-Hollandsche Uitgevers Maatschappij, 1938. Pp. 153.

The clinical fruitfulness of Ludwig Klages' procedures in graphology is due in part to the interpretation of personality around which his system of thought centers. Unfortunately, this system is self-contained to the point of self-insulation; its vocabulary, moreover, is exceedingly individualistic. It is thus highly advantageous to both the characterologist and the psychologist that some effort be made to equate Klages' conceptions to the general conceptual structure of modern psychology. It seems especially fitting that such an effort, like the present attempt, should obtain one side of the equation from depth psychology. Not only does Klages' thought share largely the philosophical atmosphere of psychoanalytical theory, but, like psychoanalysis, it maintains an isolationist attitude toward the rest of psychology.

Margadant commences with an outline of Klages' leading theses, thus affording some understanding of the basic concepts. The other eighteen brief chapters are devoted to a consideration of those properties of handwriting which Klages regards as symptomatically significant. In

each case the principles of depth psychology that seem pertinent provide the terms of the discussion.

Close familiarity with Klages' detailed work, including tabular presentations often referred to by the author, but not adequately described, would be necessary in order to decide whether the equating of Klages to depth psychology is more than an arrangement of analogues, and how much more. There are indications that the relationship is really a close one. Margadant uses as a basis for comparison the whole of depth psychology rather than some special brand of psychoanalysis. The principles utilized are of the sort called "dynamic mechanisms," such as regression, inhibition, etc., while the possibilities of Jung's typology are not neglected. If, as has been said, graphology is for Klages the psychology of handwriting, it may be that this psychology is in large measure a psychoanalysis. Although he would undoubtedly dissent, such an outcome should by no means be unexpected. Certainly, one is able easily to gain the impression from the present work that handwriting does consist of movements expressive of the personality, and that the personality is susceptible of being read, as it were, from the record, in familiar terms. The removal of the need for digesting Klages' metaphysics and the clarification of conceptions that have more or less defied outside scrutiny will alike make this book useful to all who are interested in the fields upon which it trenches. The book itself is short, but highly suggestive; it is unusually free of repetitious expositions.

Howard Davis Spoerl.

American International College.

BOOKS RECEIVED

BIDNEY, D. The psychology and ethics of Spinoza: a study in the history and logic of ideas. New Haven: Yale Univ. Press, 1940. Pp. xv+454.

CREEDY, F. Human nature writ large: a social psychologic survey and Western anthropology (with a Foreword by B. Malinowski). Chapel Hill: Univ. North Carolina Press, 1939. Pp. ii+484.

ERDÉLYI, M., & GROSSMAN, F. Dictionary of terms and expressions of industrial psychology ("psychotechnics") in German, English, French, Hungarian (with a Preface by H. L. Hollingworth). New York: Pitman, 1939. Pp. viii+98.

OJEMANN, R. H., BRANDON, V. H., GRANT, E. I., MUSGROVE, R., GABRIEL, A., & COAST, L. C. Researches in parent education IV. *Univ. Ia Stud. Child Welf.*, Vol. XVII, New Ser. No. 381. Iowa City: University, 1939. Pp. 181.

PEAR, T. H. The psychology of conversation. London: Thomas Nelson, 1939. Pp. 171.

NOTES AND NEWS

A RESEARCH grant-in-aid of $500 has been awarded by the Carnegie Corporation to Dr. Alice I. Bryan, assistant professor in the School of Library Service, Columbia University. The purpose of the grant is to enable her to work up and publish data collected during her four years of service as head of the department of psychology, School of Fine and Applied Arts, Pratt Institute, Brooklyn, New York.

A WPA GRANT to Rutgers University has made it possible to begin work on the reconstruction of an old house just off the campus for the use of the department of psychology. The partitioning into 21 rooms will furnish space for laboratory work, a clinic, offices for members of the staff, a seminar room, classrooms, and rooms for a limited number of graduate students. The interests of the department center chiefly about general and experimental psychology (Professors Sidney Sanderson and Carroll C. Pratt), abnormal and clinical psychology (Professor Griffith W. Williams), and physiology (Professor William H. Cole, head of the department of physiology and biochemistry). Off-campus clinical work is in charge of Dr. Anna S. Starr. Dr. Otto von Lauenstein, formerly of the University of Berlin, was appointed to the staff last spring, but is still unable to leave Germany. There is close coöperation with members of the staff at the New Jersey College for Women (Drs. Sidney A. Cook, head, Helen M. Richardson, and Nelson G. Hanawalt). Inquiries regarding possibilities for graduate work should be addressed to Dr. Pratt, head of the department, or to Professor Walter C. Russell, secretary of the graduate faculty, Queen's Campus, New Brunswick, New Jersey.

THE spring meeting of the Connecticut Valley Association of Psychologists on May 4 in New London will mark the formal opening of the new psychological laboratory at Connecticut College. Professor Robert S. Woodworth, of Columbia University, will be the speaker. The laboratory is equipped with a vivarium, operating room, sound-reducing room, shop, darkroom, and seminar and special research rooms.